THE JOURNAL OF PARAPSYCHOLOGY

Volume 12 DECEMBER, 1948 Number 4

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

PAGE

Editorial

The Value of Reports of Spontaneous Psi Experiences 231

Appraising Verbal Test Material in Parapsychology 236
J. G. PRATT AND WILLIAM R. BIRGE

Current Attitudes of American Neuropsychiatrists Toward Parapsychology: A Survey 257
RUSSELL G. MACROBERT

A New Method and an Experiment in Pure Telepathy 273
WILLIAM R. BIRGE

Minor Articles

An Exploratory PK Experiment Based Upon Throwing Twenty-four Dice 289
EDMOND P. GIBSON

Book Reviews

Emanuel Swedenborg by Signe Toksvig 296
C. D. BROAD

La Ricerca Psichica by Emilio Servadio 302
JAN EHRENWALD

Letters and Comments 303

Glossary 310

A SCIENTIFIC QUARTERLY DEALING WITH EXTRASENSORY PERCEPTION, THE PSYCHOKINETIC EFFECT, AND RELATED TOPICS

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

EDITORS

J. B. RHINE, J. G. PRATT, and BETTY M. HUMPHREY, *Editors*
J. A. GREENWOOD and T. N. E. GREVILLE, *Statistical Editors*
DOROTHY H. POPE, *Managing Editor*

ADVISORY EDITORIAL BOARD

HORNELL HART
Duke University

H. H. PRICE
University of Oxford

GARDNER MURPHY
College of the City of New York

ROBERT H. THOULESS
University of Cambridge

This JOURNAL is published on the fifteenth day of March, June, September, and December by Duke University Press, Durham, North Carolina.

Contributions submitted for publication and all editorial communications should be addressed to the Managing Editor, Dorothy H. Pope, College Station, Durham, North Carolina. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, the JOURNAL cannot guarantee that they will not be lost in transit, and contributors are urged to keep copies of their papers. All contributions should be typewritten, double spaced. References should be given in the form adopted by the JOURNAL.

Reprints may be ordered when the proof is returned. The JOURNAL will bear one half the cost of reprints up to two hundred copies. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to Dorothy H. Pope, College Station, Durham, N. C.

The subscription price is $5.00 a year; single current numbers $1.25. The rate for back volumes is $6.00; for single numbers $1.50. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. All remittances should be made payable to the JOURNAL OF PARAPSYCHOLOGY.

Entered as second class matter May 21, 1942, at the post office at Durham, N. C., under the act of March 3, 1879.

DUKE UNIVERSITY PRESS

The Journal *of*
Parapsychology

Volume 12 DECEMBER, 1948 Number 4

EDITORIAL

THE VALUE OF REPORTS OF SPONTANEOUS PSI EXPERIENCES

RESEARCH in parapsychology, as in most fields, had its roots in natural events which occurred nonexperimentally, that is, spontaneously. Reports of phenomena of thought transference and clairvoyance stimulated the earlier explorers to look for experimental evidence of these capacities, and the character of the phenomena determined the kind of experiments that followed. For example, if a case of telepathy occurred in a demonstration of hypnosis, subsequent experiments would naturally involve the hypnotic state, at least for a time. Or, if the spontaneous phenomenon seemed to involve clairvoyance, the experiments were based on clairvoyance tests. Even the individual subjects used in the early experiments were discovered through their nonexperimental experience of the abilities in question.

The value of these spontaneous cases obviously lay in the start and direction they gave to experimental investigation, and the importance of these functions can hardly be exaggerated. The cases did not, however, in themselves constitute adequate evidence of any form of extrasensory perception; if they had, further experimental work would not have been needed. As it was, the conclusions eventually reached by scientists regarding the occurrence of telepathy or clairvoyance rested entirely on evidence gained from the experiments to which the case material had been a stimulus.

Unfortunately the importance of this restriction of the role of spontaneous cases has not been generally recognized, and therefore this material has often been miscast as evidence of the abilities involved. Because of the elusive nature of ESP phenomena, it has not been easy to follow up experiences with experiments which would produce evidential results on demand. Accordingly there has been a great deal of effort to recruit and dress up these spontaneous experiences themselves as evidence of reliable character, but all attempts to authenticate the reports well enough to allow reliable conclusions to be drawn have been comparatively unsuccessful. The result of this abuse of the case report has been a misunderstanding of its true significance and usefulness. In consequence the value placed on case reports in parapsychology underwent a decline, and they have been pretty well ignored in recent decades.

If, however, the spontaneous experience can be given its proper role, it can play a great part in the researches ahead, perhaps a crucially important one. Such material can help us to initiate new experimental approaches, not any longer to establish psi, but to find out more about what it is and how it operates. And once we abandon all interest in using these experiences as evidence of psi capacities, we have available a wealth of case data which would otherwise be ignored. If we use case reports only to give us suggestions for new experimental projects, it becomes relatively unimportant whether or not supporting affidavits are available in the record, whether any witnesses were present, or what the education or standing of the percipient or reporter. In any right view of the matter, a spontaneous case can serve to contribute a hypothesis, raise a question, or furnish a possible insight, but it cannot go far beyond this point. Since all these hypotheses, questions, and insights will have to be subjected to crucial experimental tests before any conclusion is reached, the demands of authentication that have been the bugbear of case studies in parapsychology are seen at once to be largely irrelevant.

The worst thing then that could possibly happen for giving undue consideration to an unsound case report would be that it would lead to an experiment that would yield no confirmatory results. The odds are very strong, however, against any single case ever leading

to an actual experiment. No experimenter is likely to be carried so far by an unsupported case. He will, however, be led to look for more cases of the same general type, and, if he finds enough others to support it in the particular suggestion it makes regarding the nature of psi, he would be warranted in conducting a crucial test of the hypotheses. We may expect a great deal of extraneous and erroneous material in any case collection, but the worthless and misleading elements will likely not compound in the course of the analyses. They are likely rather to be screened out, while the true characteristics will recur from case to case, sufficiently for our purposes of acquiring new leads to better researches.

In the early case studies (up until the question of spirit agency came into dominance) the primary interest in mind was the question of the occurrence of ESP, mainly of telepathy. But, as always happens, the experiments restricted the natural function under investigation, and the subsequent research developments became increasingly narrowed and specialized as the controls were increased. In time the experimenters got far away from the original phenomena and may have lost a great deal of what was represented in the pre-experimental manifestations.

Ideally the types of spontaneous experiences that initiated experimentation ought to have been kept in perspective throughout the course of the research. As it is, they have been neglected, and we must now go back to them again for a fresh outlook. And while we are no longer concerned with the original problem of psi occurrence, the very same types of experiences that led to the original experiments in telepathy and clairvoyance may now be of service again in giving us clues as to the nature of these capacities. We cannot, in fact, conceive of a richer source of suggestions as to the way in which the psi processes may best be approached and studied. We may well count upon the ability to give itself away more revealingly in its nonexperimental manifestations than in anything which we could hope at this stage to demonstrate in the laboratory.

Accordingly, it is one of the most urgent needs of our research field that we turn back to these natural springs of research ideas and draw upon them to the fullest extent. The realization of the importance of this source material has for some time motivated a

study of cases at the Parapsychology Laboratory at Duke, and reports of this study will be forthcoming in due time. The general plan of the program is now fairly well drawn up; the first step has already been taken, that of setting up certain general categories representing the major problems confronting us in parapsychology today. Now, as the reports of psi experiences come in, they are classified and indexed under one or more of these categories according to their bearing. (Only the obviously worthless reports are rejected.) After a given collection is large enough to justify study, those who are concerned with a particular research problem can find already assembled for them such case material as bears upon their special interest.

In going to work upon this material, each research worker will, of course, wish to classify the cases in his category still further, using headings related to his own scientific needs. On every main problem being dealt with there are many subordinate questions, and a further analysis will have to be made to break down the material of a given category into subdivisions related to these questions. The result may be that, let us say, a group of fifty cases are available that bear on a given question confronting the individual research worker, with perhaps half that number (or fewer) actually contributing meaningful material. From a study of these cases one can count on getting at least a little new direction for his thinking, and perhaps a genuine reorientation toward his problem. In any event he cannot be seriously misled.

There are ways, too, of making this case analysis method a still more fruitful one. For example, if a number of large collections can be made and analyzed on similar lines, the results from one collection may be used to check against those of another. If a research suggested on the basis of one collection gets support from another one consisting of independent cases, it is in a much stronger position indeed. Any project that would involve heavy research effort and expenditure might have special need of such support.

One further point is evident. The job of making adequately large collections and of sifting, sorting, and breaking down the material into usable form is an enormous undertaking. It is too much for one laboratory, as such laboratories are staffed today. It

is the kind of project that requires large resources that would permit it to be farmed out for the cooperation of a number of qualified collaborators. We dare not wait, of course, for such an ideal situation before we begin. In fact everyone capable of any sort of activity in parapsychological work can be of help in the collecting and forwarding of spontaneous cases as he may encounter them to the laboratories where they will be preserved and studied. Every reader who appreciates the importance of such case reports can himself become a collection agent for this source material on which the guiding insights for the research of the future may greatly depend.

J. B. R.

APPRAISING VERBAL TEST MATERIAL IN PARAPSYCHOLOGY

By J. G. Pratt and William R. Birge

ABSTRACT: The writers trace the efforts that have been made toward finding a satisfactory means of assessing the accuracy of verbal test responses. The descriptive statements which make up this material appear in many forms: as mediumistic utterances, as impressions given in association with a personal object, as automatic writing, and the like. Progress toward a method is described step by step, beginning with the work of Saltmarsh in 1930. The emphasis is placed upon the advance represented by each new contribution. The current procedure is explained, and suggestions are made for further research.—Ed.

Introduction

This will chiefly be an account of steps that have been taken in the Parapsychology Laboratory over a period of fifteen years toward the development of a method of evaluating verbal material obtained in tests of psi abilities. It will be a report on a search for a method, a search that has resulted in definite progress even though it is still not complete. Thus far in working with responses given in the form of descriptive statements, the investigators have been trying mainly to find a way of appraising the results, and relatively little attention has been paid to obtaining the best possible subjects for such psi tests and to providing the most favorable psychological conditions.

Verbal material is one of the forms in which ESP has been most frequently reported to occur. If an adequate method can be found, a careful study of this type of response obtained under suitable test conditions would obviously have wide application and great value. The type of word material that has received the most attention in parapsychology has been the utterances of "mediums." These statements have been studied chiefly from the point of view of their bearing upon the hypothesis of spirit agency. The study of mediumship depends largely on a method for accurately evaluating verbal material, and we may say that progress in such study has long awaited the development of methods such as we are seeking.

Early Efforts at Appraisal

In spite of the long history of the investigation of mediumistic records, the first efforts to find a way of appraising verbal material were made less than two decades ago. The pioneer in this line was H. F. Saltmarsh, who published the first report on this topic in 1930 (3).

Saltmarsh used as the subject of his experiments the English medium, Mrs. Warren Elliott. Sometimes the person for whom the utterances were intended—that is, the "cooperator" in a test—was present at the time the subject gave her responses, but at other times the tests were made with the cooperator absent. In the latter case, the subject was handed some small personal belonging and was asked to give her impressions regarding the circumstances of the people who had been closely associated with the object. It was these "token object" tests that Saltmarsh used in trying out new methods of evaluation.

In the method of scoring that Saltmarsh first proposed, the procedure was as follows: To start with, he broke the subject's verbal material down into separate items or statements. He then submitted each of these itemized records to the cooperator to be annotated regarding the applicability of the statements to either his own personal circumstances or those of the deceased owner of the token object. At the same time, he submitted each record to a "control" cooperator, asking him to comment on the items as if the information were intended for him. On the basis of the comments made, Saltmarsh decided whether each item was true or false for the cooperator and for the control annotator.

This procedure for marking the records was a distinct advance over anything done up to that time, but as a pioneering effort it naturally had some weaknesses. The most serious of these was the fact that the cooperators knew they were commenting upon material that was intended for them, and there was no way of knowing how this knowledge affected their scoring of the items. Also, the control markers knew that the information was *not* directed toward them, and this fact introduces a real doubt as to whether their judgments were comparable with those of the cooperators. Finally, Saltmarsh himself knew which comments were made by the cooperators and

which ones by the control markers. It is possible that his judgments reflected some degree of subjective bias in deciding in each instance whether an item was true or false.

The method involved several further steps before a final assessment of the material was reached. Saltmarsh classified the items of each record in terms of three groups: clichés, which would be widely applicable; definite statements, which would apply less widely; and characteristic statements, which would be true of very few people. To these he assigned values of 1, 5, and 20, respectively, for each item, and on the basis of these values he then gave two scores to each of the subject's test records, one for the cooperator and another for the control annotator. The criterion of significance arbitrarily set up was that the cooperator's score on his own record should be at least eight times as great as the score obtained by the control marker.

In the second attempt to improve the methods of appraising verbal test material Saltmarsh was assisted by S. G. Soal (**4**). This study made use of a more precise statistical procedure for measuring the significance of the results. The basic mathematical method was suggested by Soal and further improved by R. A. Fisher. It provided a means of evaluating a series of items regarding each of which two things were "known": (1) whether each was *true* or *false* for the actual cooperator; (2) the value of its probability ratio, or its likelihood of being true of anyone selected at random from Western adult civilization. (Saltmarsh arrived at the "truth" or "falsity" of those statements that were relevant to each cooperator, as he had earlier, by having each one annotate his own material and then by deciding himself whether or not the item applied. He also assigned the item probability ratios by personal judgment.) Given these two sets of facts, the method made possible an assessment of the results in terms of the probability that the results would by pure chance have fitted the cooperators as well as they did.

As we see it now, there was a fundamental weakness in this proposed statistical analysis that was caused by the nature of the material to which the method was to be applied. The Saltmarsh-Soal formulas assumed that the items to be evaluated, or the data derived from them, were independent of one another. There is no

way of knowing that this assumption is justified when the items are drawn from general descriptive accounts. For example, the subject in a token object test of his psi abilities might start out to talk about an English cottage. Having chosen this theme, he might go on to give several items appropriate to it. If an English cottage happened to have some personal significance for the cooperator involved, many of the items might be correct simply because the description was internally consistent. On the other hand, if the general topic of the description happened to be wrong, all the items might be rejected for this reason. Either way, the items would not yield statistically independent data for analysis.

Quite apart from the interdependence of the items themselves, this difficulty in the way of statistical assessment might arise from subjective factors of judgment on the part of those who score the items. If a person knows that the information he is asked to check is intended to fit his circumstances, he may score it differently than he would if he thought that it was not his. This emphasizes the importance of having the cooperators ignorant of whether or not the material they are asked to mark is their own.

A further weakness in the new method as it was first applied lay in the fact that Saltmarsh arbitrarily decided the item probability ratios on which the evaluation was based. That is, he used subjective judgment in deciding how likely it was that a given item would be true purely by chance. In making a trial application of the method to one of Mrs. Elliott's records, he pointed out that it would have been better if the average of several independent judgments had been used.

The work of J. F. Thomas (5) represented the first large-scale attempt to apply the Saltmarsh-Soal method of scoring verbal material. Making the advance in methodology that the earlier workers had suggested, he averaged the opinions of several judges to get the probability value of each item. He was the person (cooperator) for whom all the test statements were intended, and he scored them all as being correct, incorrect, inconclusive, or unverifiable. Only the first two classes, the "true" and "false" items, entered into the final analysis. Two separate committees of judges were asked to assign probability estimates for these items, and duplicate evaluations

were worked out on the basis of these two sets of ratios. In one case, a critical ratio of 13.04 was obtained; in the other, 4.91. Each of these is significant, but the wide discrepancy between them suggested to Thomas a weakness in the method of using estimated probability ratios.

A second statistical treatment applied to these records was that of having a number of "control" cooperators each check the material for applicability to himself, just as Thomas had done. By an appropriate method of analysis Thomas' score was then compared with the distribution of scores made by the control group and a measure of its significance in terms of this general background was obtained.

Thomas' methods of assessing the subject's success in giving correct information suffered throughout from the same basic weaknesses as the earlier efforts. He did not achieve objective judgments of the material, as everyone who scored it knew whether or not it was intended for him. Also, the use of the Saltmarsh-Soal formulas here, as in the original study, involved the unwarranted assumption that the items within each record, or the data derived from the items, were statistically independent.

The first efforts to advance the methods of appraising verbal test material made in the Parapsychology Laboratory were based upon experiments in which Mrs. Eileen Garrett was the subject. There were two separate series of tests, one in the spring of 1934 and the other a year later, and a full report of this work was published (2). In the present paper we shall be concerned only with reviewing briefly the advance in methodology that each series represented.

In the first series of tests with Mrs. Garrett a new method of deriving the item probability ratios for use in the Saltmarsh-Soal formulas was introduced. In this experiment there were twelve cooperators, and one session was held for each, with the cooperators present. The twelve records were later broken down into separate items by the experimenter and all of them were submitted to all of the cooperators, who were asked to check each item as correct (√), incorrect (X), unknown (?), or not relevant (in this case the space within the parentheses was to be left blank). Only those items from each record which were marked either as correct or incorrect

by the actual cooperator were used in the final evaluation. The probability ratio of each of these items was obtained from the way the other eleven cooperators checked each one. The denominator of the ratio was taken as the combined number of those marking the item as correct or incorrect, while the numerator was the total of those checking it as correct. When the actual cooperator was the only one who checked an item as correct, his marking of the item was included with the other eleven ratings in getting the probability ratio. This was done to avoid having an item probability of zero, which would occur when the statement was true for no one except the actual cooperator.

The advance represented by the first series, therefore, was that of having the cooperators themselves provide the control checking of the items. In this way it provided empirical item probability ratios and met the objection offered by Thomas to the earlier attempts to use the Saltmarsh-Soal method.

The second series of tests with Mrs. Garrett made a still further advance toward an adequate method of assessing verbal material. It introduced conditions to keep the cooperators unaware of which were their own records until after they had marked them all. Fifteen cooperators took part in this experiment. At the time of each test, the cooperator concerned was kept in a room adjoining that of the subject, with the intervening door closed. The identity of the cooperators was withheld from the subject, who worked in this series, as she had in the former one, in a trance state. Each cooperator furnished a token object. The subject knew that these were available, but she did not always use them.

After the fifteen tests had been conducted, the experimenter analyzed each record into items or points as in the first series. As each one was prepared it was mailed out to the cooperators to be checked. Each cooperator was asked to mark all the material as if it applied to himself. Since none of the cooperators was in the room with the subject at any of the tests, they were not able to identify their own records from memory. This control of the subjective factor that was present when the cooperator knew which record was his own was the outstanding advance in method achieved in this series of tests.

From the scored records, the item probability ratios were obtained as they had been in the first series. These were then substituted in the Saltmarsh-Soal formulas to arrive at an evaluation of the results. The evaluation showed that three of the fifteen subjects who took part in the second series obtained an individually significant rating on their own records, and that the fifteen records as a whole were highly significant.

Was this an adequate method for testing the significance of these results? The original report discussed one possible loophole in the procedure. Some of the cooperators might have obtained significant scores on their individual records because they consistently accepted a large percentage of items throughout the entire fifteen records. To check up on this possibility, a study was made of the "yes"-tendency in the three cooperators whose records turned out to be significant. In two instances there was no general tendency to accept a large percentage of items in all the records, but there was in the third case. In other words, the results of two significant tests were not attributable to this factor. It was pointed out, however, that even such a general "yes"-tendency as was exhibited by the third cooperator would have the effect of *lowering* the significance of the other records and thus the exaggerated value given his own material would be offset to some degree when the series was evaluated as a whole.

However, there is still a serious difficulty that prevents our concluding that the subject demonstrated psi ability in these tests. This is the matter of the interdependence of the utterances given in each session, the characteristic of descriptive statements that was explained in connection with the first application of the Saltmarsh-Soal method. It is a weakness that appears to be inherent in the method itself, a fact that has not always been as fully appreciated as it is today. Finding some means of overcoming this difficulty has become one of the primary objectives of the efforts to devise a method of assessing connected statements.

Methods Tried More Recently

Since the publication of the preliminary report on the assessment of mediumistic utterances in 1936, the Parapsychology Laboratory has continued the efforts to improve the methods for dealing with

verbal test material. During this period we have worked almost exclusively with the results obtained in token object tests. We used a number of subjects who thought they might be able to demonstrate their psi abilities in this way. The objects were always securely wrapped and sealed to avoid the possibility of inferring anything from them. In most instances, the cooperators submitted objects that were keepsakes of high sentimental value because of their connection with someone deceased. In this way we were able to work more effectively with subjects who in their experiences along these lines had been oriented toward the spiritistic hypothesis.

As we stated earlier, the emphasis thus far has been kept on the development of a suitable way of appraising verbal test material, and we shall continue to limit ourselves in this paper to this question. The results, which have not been significant as a whole, will be presented only as needed to illustrate a method. Nor shall we be concerned at this time with the details of procedure in these token object tests that do not directly affect the assessment technique. Suffice it to say that precautions have been taken to insure that the subjects should not obtain any clues regarding the identity of the cooperators and the original owners of the token objects.

In addition to the shortcomings discussed earlier, all the methods already described suffered from the practical difficulty of being cumbersome to apply. This difficulty became especially great in the methods that were first tried in the Parapsychology Laboratory. The use of control checking of the material by the cooperators led to the necessity of treating a larger number of test records at a time in order to increase the reliability of the item probability ratios. As work along these lines was resumed, it was with a clearer recognition of the fact that the method of assessment would have to be one that investigators would find practicable.

Another practical difficulty was that some of the methods had been wasteful of the subject's utterances themselves. A literal interpretation of the statements had been followed, and as a consequence many of the items were excluded from the final evaluation as having no relevance for the cooperator for whom they were intended.

In 1944 C. E. Stuart proposed a method that seemed to overcome these two practical difficulties to some degree. One suggestion

he made was that when the records were broken down into items the wording used should be such that each cooperator would be able to mark a larger percentage of the statements as either correct or incorrect for himself. Stuart took the view that the subject might be giving correct information in his utterances, but that he might be directing his remarks less accurately than a literal interpretation would suggest. Such things as mistaken identity or the mixing up of two or more individuals connected with the cooperator might be occurring. Thus, suppose the subject said of a particular cooperator that he had a deceased uncle and that this uncle had been fond of smoking a corn-cob pipe. According to the earlier methods of itemizing the material, the descriptive item about the corn-cob pipe would necessarily be considered of no relevance to the cooperator if he did not have a deceased uncle. With Stuart's method, the item might still be considered applicable provided the cooperator had a deceased father or some other relative who fitted the description. For example, this particular item might be presented in two parts, in some such manner as follows: "You have a deceased relative who used to be fond of smoking a corn-cob pipe ( )." "This relative was your uncle ( )." This broader basis for the interpretation of the material proposed by Stuart was in keeping with the view that the subject's verbal responses might be even more "free" than the actual words of the subject seemed to imply.

Stuart also suggested a method of evaluation that required fewer cooperators for an experiment. He proposed five token object tests as the standard number. An ingenious evaluative procedure was devised that still made use of empirical ratios but did so with only five cooperators. As in the earlier work, each cooperator was asked to score all the records without knowing which one was his own. Stuart selected for evaluation only those items to which four co-operators had responded with "no" and one had responded with "yes." On a purely chance basis, each of these items has a one-fifth probability that the person who scored the item "yes" would be the one for whom it was intended. If the cooperator who scored the the item "yes" was the same cooperator for whom the subject intended that item, it was counted a hit. Otherwise it was a miss. The statistical significance of a particular record as well as of a series

of five records could then be computed in exactly the same manner as that in which the deviation for a particular number of trials in a standard ESP card test is measured.

Stuart had not made a final formulation of his method before his untimely death in March of 1947. His procedure as described here offered definite advantages of the sort that he was trying to achieve. There were still some shortcomings in the procedure, however, as Stuart himself was well aware. One of these was the fact that the method, in spite of achieving a much higher percentage of judgments of the material from the cooperators, still allowed only a small amount of the information given by the subject to enter into the final calculation of the results. Another objection, one of an even more fundamental character, was that the statistical evaluation used still involved the assumption that the items were independent of one another. This requirement concerning item independence was not met in this procedure any more than it had been in the methods already described.

Stuart also made an effort to apply to verbal material the evaluative technique which he had developed for free response tests based upon pictures as targets. This is the preferential matching method that is now widely used in drawings tests. There seemed to be no a priori reason why this method should not be used for free verbal material.

In applying this method, Stuart broke down each record in a set of five token object readings into separate paragraphs dealing with distinct topics or personalities. The records were then sent to the cooperators in units of five paragraphs, including one paragraph from each of the original five records. The cooperators were asked to rank each paragraph in a unit according to its applicability to their circumstances. This method was later discarded as too wasteful, since it lumped together all the information in a paragraph as a single trial by the subject.

After Stuart's death, one of the writers (W.R.B.) for a time assumed an active role in the effort to improve the methods for handling free verbal material. Still another new approach was tried. As before, the information in a set of five records was itemized and submitted to the five cooperators without their knowing which were

their own records. For the evaluation of the results, only those items in all the material that were checked by each cooperator were counted as "trials" made by him. Of these trials, those check marks which came within his own record were considered as hits. If the records were of the same length, in the sense that each one offered the same number of items to be accepted or rejected, the expectation was that the number of check marks made by a cooperator within his *own* record would be one-fifth of the total number of his check marks in *all five* records. Any observed tendency for a cooperator to check a larger number of items correct within his own record could be evaluated in terms of the deviation from the expected number by the same formula as Stuart had used—the one that is commonly applied in ESP card work.

This method did seem to represent a definite step forward. It was a more simple and direct approach to the problem than the procedure Stuart had proposed, and it also enabled the investigator to make use of all the items that the cooperators checked as correct for themselves. Its shortcomings were that: (1) like the earlier procedures, it assumed an independence of items; (2) records of different lengths presented a special difficulty in that they could not be used without changing the 1/5 probability basis of evaluation required for the statistical analysis. The practice proposed was that of itemizing all the records and then dropping off items at the end so that all the records would have the same number of items as the shortest one. This device again involved the wasteful necessity of leaving out some of the material.

The Current Procedure

Up until this time all of the methods that had been tried, except that of treating paragraphs as a whole by means of the preferential matching method, involved editing of the free verbal material, or breaking it up into items before it was scored by the cooperators. The practice in most of the methods of assessment tried had been to separate out the items and to do some paraphrasing of the subject's utterances in order to help the cooperators in their marking of the material. The idea occurred to one of us (J.G.P.) that keeping the exact words used by the subject would offer several distinct advantages. In the first place, this procedure would save a great

deal of time. Furthermore, it would avoid the danger of misinterpreting the verbal material in a manner that might affect the scoring. This danger, of course, becomes serious only in case the person who itemizes the material is acquainted with the cooperators. It seemed that it might be best to avoid, if possible, all "editing" of the records. Experience with these records suggested that it might be possible to achieve all the advantages of itemization simply by inserting parentheses ( ) wherever the subject made a remark that introduced a new thought or qualified a statement in any way. The only judgment required was in deciding where to insert one of these checking points in the material.

All of the foregoing methods for evaluating verbal test material in terms of separate items assumed an independence among the items that may not have existed. For the items within a record to be truly independent, the fact that a cooperator has checked a particular item as correct or incorrect should have no relationship to the manner in which he checks the remaining items in the record. There are two main reasons why the items cannot be considered independent in this sense. First of all, descriptive accounts tend toward some degree of self-consistency: items pertaining to the same topic are likely to be highly related. It follows that, if a particular item is correct or incorrect for a cooperator, an indeterminate number of other items are likely to be checked as correct or incorrect purely as a consequence of this interrelationship or self-consistency.

The second factor undermining the assumption of independence of items is the fact that, quite apart from the self-consistency of a record, the answers made by the cooperators may not be independent of one another. Thus a cooperator may form the general opinion that a record is or is not meant for him on the basis of a few introductory items. Through a perseverative tendency, the cooperator's scoring of the remaining items in the record may show a general effect of this opinion.

The method currently being tried out in the Parapsychology Laboratory does not assume independence of items. The logic of this new method may be explained by an illustration drawn from familiar areas of ESP research. Investigators have long been aware that data obtained from group tests may not be interpretable

in terms of the same statistical analyses as apply to the data of individual tests. In group tests, a number of subjects attempt to call the same targets, and there is no way to be sure that any non-randomness found in the calls made by different members of the group could be attributed only to the special ability (ESP) being tested. For example, subjects in a group might exhibit a tendency to start their calls with a certain symbol, or they might show the same symbol preference in their series of calls because of cultural or environmental influences or for some other reason having nothing to do with ESP ability. Any such group pattern or characteristic of response would not change the expected average number of hits, but it would affect the variance, the distribution of total scores about the mean for the group. The data could not properly be evaluated on the binomial hypothesis, which assumes that the calls are independent. For the variance of the binomial hypothesis to apply, it is necessary that at least one of the series of events to be compared—either the cards or the calls—be random. This condition holds in a test in which a single subject calls a particular card order only once; for the test to continue, a new card order is provided. In this case the use of a new order of target symbols for each run meets the requirement for randomness. The difficulty in the group situation is that the same card order is used over and over again in checking the calls of the individual subjects.

What statistical procedure might be used to evaluate group tests without making the assumption of independence of results when many call sequences are compared with the same card order? Greville worked out and published a solution to this problem (1). His procedure involves taking the actual distribution of calls in the separate trial positions (the calls actually made on each target symbol) as the given data. For example, consider the case of 100 subjects attempting to call through a single random order of 25 ESP symbols. The results would be evaluated in terms of 25 trials, not as 2,500, the actual number of calls made. The statistical question asked is as follows: Given the particular distribution of symbols in the calls as observed in each of the 25 trial positions, what is the probability that this fixed distribution would give a total score as high as the one made by the group? The total score on the test

would be found by comparing the distribution of calls with the order of symbols in the target deck, this order being a random arrangement which is only one of a large number of possible permutations of the 25 symbols. For each of these permutations of the target deck, a score could be obtained by checking the order of symbols against the "fixed" or observed distribution of the subjects' calls. If the deck were arranged in every possible way and the scores were checked for every permutation, it would be possible in theory, though impracticable in fact, to work out the mean score and the variance of the distribution of all these scores. The score made by the group in the actual test (that is, the hits obtained when the observed distribution of calls is checked against the particular random order of 25 card symbols used) could then be measured in terms of its deviation from the mean and the standard deviation of the entire distribution.

Greville developed the formulas for applying this particular test of significance without the necessity of actually permuting the target order through all its possible arrangements. He considered two general situations: first, that in which the target order represents a truly random selection from among the choice possibilities offered; secondly, that in which a closed deck is used, presenting an equal number of all the possibilities with a random arrangement such as might be given by adequate shuffling. Only the closed deck situation is relevant to the problem of evaluating verbal material.

In thinking over the difficulty which the interdependence of items in long descriptive accounts seemed to cause for statistical assessment, it occurred to one of the writers (J.G.P.) that the Greville method for evaluating the data of group card tests could also be applied to the results obtained from token object tests. As a means of making this application easier to follow, we shall first give another illustration of the use of the method in evaluating the results of a group test involving card calling. In this instance we shall set up a hypothetical situation that is more closely parallel to that represented by token object tests.

Assume that a group of 100 subjects is being tested for ESP ability. They are told that the test will consist of five trials. For targets, cards bearing the common surnames Jones, Brown, Smith,

Hill, and Greene will be used, each target being used only one time during the five trials. The target order is thus one of the 120 possible permutations of the five target names.

Assume, further, that the instructions are given that each subject is to make a response on any particular trial only if he feels confident of making a hit. This would lead to differences in the numbers of calls from trial to trial, a condition which presents no difficulty in the use of the Greville method of analysis. The method, in other words, makes possible the computation of the mean and the variance of the scores that make up the general distribution to which the actual score belongs regardless of unequal numbers of calls from trial to trial. Those who are interested in the technical statistical aspects of this method will see the reason for this from Greville's original article. Others may accept it as a statement of fact.

Assume that when the data are tabulated in terms of the frequency of calling each name on each trial, the following distribution of responses on the five trials is found:

Trial	Subjects' Responses				
	Jones	Brown	Smith	Hill	Greene
First	9	7	0	0	3
Second	3	0	0	0	12
Third	10	0	26	3	13
Fourth	2	0	0	2	12
Fifth	3	0	0	4	4

The Greville method takes these figures as the given data, and it thus avoids making any assumptions regarding statistical independence among these observed responses. Taking these data as they are, the Greville method enables us to find the *mean* and the *variance* of the 120 scores that would be obtained when this particular matrix of responses is checked against the 120 different orders in which the five target names might have been presented.

The reason it does not matter whether these calls were independent or not is simply that the method makes no assumptions regarding how they came to be distributed as they are. The analysis merely assumes that the order in which the five stimulus names are presented is random, and this requirement is fully met by the conditions. For example, the names might have been those of five

persons known to the group of subjects, and the wide variation in the number of names called might have reflected differences in popularity. The Greville test of significance would still give a correct probability figure showing how frequently, on a purely random basis, this highly biased set of responses on the five trials would correspond with the order of target names as well as was found to be true in the particular instance observed. If the probability figure meets the accepted criterion for statistical significance, the indication of a causal relation between the responses and the random order of target names is precisely as strong as in an experiment analyzed by any other method if the two sets of results happened to be significant at the same level.

Suppose that in this instance the random target order Brown, Hill, Smith, Greene, and Jones had been used. By summing the figures from the appropriate column for each trial, we find that the score of the group was 48. To determine whether or not this is significant, we only need to derive the mean and the variance by the Greville statistic, find the deviation of this score from the mean, and then divide the deviation by the standard deviation (square root of the variance) to arrive at a critical ratio for the test. All that this test of significance assumes is that the target order used was selected at random.

Keeping this illustration in mind, we may turn now to the question of how the Greville method of evaluation can be applied to verbal material obtained in token object tests. An example of its application to the records actually obtained in a set of five such tests will be given. For convenience, let us designate the cooperators as Jones, Brown, Smith, Hill, and Greene.

When the five records of this set were received from the subject, they were prepared for marking by the insertion of checking points within the verbatim statements. The five records were then arranged in random order and each one was given a code designation. They were then typed with sufficient copies and all five records were sent to each of the five cooperators to be checked throughout. The cooperators were given instructions to mark the items within each record on the assumption that they were all intended for them, and to use a check mark ($\surd$) to show a correct

statement, a cross (X) to show an incorrect one, and (?) for a doubtful item. In general, the instructions were so worded as to encourage a liberal interpretation rather than a restricted one. The aim was to get the widest possible measure of application of the material to the personal circumstances of each cooperator.

After the marked records were returned by all five of the co-operators, the items checked as correct were tabulated for each one, thereby showing the number of items he marked as correct in each record. The following distribution of check marks was found:

Records	Cooperators' Check Marks				
	Jones	Brown	Smith	Hill	Greene
First	9	7	0	0	3
Second	3	0	0	0	12
Third	10	0	26	3	13
Fourth	2	0	0	2	12
Fifth	3	0	0	4	4

These figures mean that Jones checked nine items as correct for his circumstances in the first record, three in the second, etc. These figures, reading them horizontally, could be thought of as meaning that the first record was called "Jones" nine times, "Brown" seven times, etc. Thus the figures take on the same significance as the distribution of subjects' calls in the illustration of the group test with five cards bearing the same names as shown on page 250 above. In fact, we have used identical figures in both illustrations to emphasize the close similarity of the two situations from a statistical point of view.

In the case of the token object test illustrated here, the "owners" of the five records were the targets. These were presented in one of the 120 possible permutations, an order that was selected at random and kept secret from the cooperators until after they had done their checking. It is this random order of "targets" that provides the basis for applying the Greville statistic, and the evaluation gives a straightforward statement of the probability value of the observed score obtained from the way in which the cooperators distributed their check marks.

In the example we have presented, the first record was Brown's, the second Hill's, the third Smith's, the fourth Greene's, and the fifth Jones's. The total score and the evaluation are exactly the

same as for the group card test with the random order of target cards assumed in that instance.

If significant results were obtained on verbal material appraised in this manner, the interpretation would be made in the same manner as for any test of significance. The method would permit the exclusion of chance as a reasonable explanation of the results, and this is all that any statistical analysis can do. If only five records were used in an analysis, it would not be possible to obtain a P-value of less than $\frac{1}{120}$ for a single set. When small sets are used, the results of a number of them combined might form the basis for any conclusion.

The reader who is interested in the more technical aspects of this method is referred to Greville's original article. In the appendix we have presented the basic formulas only as far as they are necessary for the evaluation of verbal material. Also, the complete evaluation of the data from the set of token object tests described in this section is presented in the appendix.

Suggestions

The following appear to us to be the most urgent research needs:

1. It is important to find out how sensitive the Greville method is when using different numbers of records in each analysis. As some of our colleagues have suggested, sets of five tests may be too small for the best results. But for convenience and speed in handling the results, small sets are advantageous. Therefore the aim should be to keep the sets as small as possible without making the test of significance too insensitive.

2. For the purpose of further refining the statistical practices, it is important to have material that will give significant results. Ideally, this should be obtained from actual experiments with psi capacities. But if successful subjects for token object tests are not available, verbal records might be deliberately made up to fit certain cooperators in order to see what size set is to be preferred. This

method has already been used with good effect in the Parapsychology Laboratory in comparing different methods of evaluation.

3. The psychological conditions, too, are important. The results obtained thus far suggest that in this respect the tests have not provided the essential requirements for success. For example, the use of sealed token objects may be an unfavorable way of working. It might be better to use exposed token objects that would not reveal anything regarding their owners, such as similar buttons or keys. Heretofore we have been sending all the token objects for a set of tests to a subject at one time, and there have been some indications that under these conditions a confusion among the cooperators comparable to the displacement effect may have occurred on several occasions. The results suggest that it might be preferable to send out only one token object at a time. Both patience and ingenuity are needed to devise test procedures that will satisfy the statistical requirements and provide favorable psychological conditions at the same time.

Appendix

The use of the Greville method in evaluating verbal material may be illustrated by the data from the set of token object tests given in the paper. We shall present the formulas in convenient computational form without going into technical questions of derivation and proof.

For the analysis, a slightly different arrangement of the figures from that shown on page 252 will be convenient. The columns showing the number of items the cooperators said were correct for themselves may be arranged across the page in the same order as the cooperators down the left-hand margin. The marginal totals of the figures within the matrix are needed, as well as the total number of times items were marked as correct. The data as prepared for evaluation are then as follows:

Cooperators	Cooperators' Check Marks					Total
	Brown	Hill	Smith	Greene	Jones	
Brown.....	7	0	0	3	9	19
Hill.....	0	0	0	12	3	15
Smith.....	0	3	26	13	10	52
Greene.....	0	2	0	12	2	16
Jones.....	0	4	0	4	3	11
Total.....	7	9	26	44	27	113

The number of items checked, or the sum of the figures in all 25 cells of the table, constitute the "calls." In this case, the number of calls is 113.

The number of "hits" is the sum of the figures on the main diagonal, or 48.

The mean number of hits expected is one-fifth of the number of calls, or 22.6.

This particular set thus gave 48 hits where 22.6 were expected, or a deviation of +25.4.

In order to compute the variance of the scores from the observed distribution, the following values are required:

The square of the number of calls (N^2) = 12,769

The sum of the squares of the individual cells (Σa^2) = 1,439

The sum of the squares of the separate row totals (Σr^2) = 3,667

The sum of the squares of the separate column totals (Σc^2) = 3,471

If n is the number of token object tests in a set, the variance is given by the expression:

$$V = \frac{1}{n^2(n-1)} [N^2 + n^2 (\Sigma a^2) - n(\Sigma r^2) - n(\Sigma c^2)]$$

For the present case we have:

$$\frac{1}{25(5-1)} [12{,}769 + 25\ (1{,}439) - 5\ (3{,}667) - 5\ (3{,}471)] = 130.54$$

The SD = $\sqrt{V}$ = 11.43

The CR of this set is therefore $\frac{+25.4}{11.34} = +2.22$

(P = .013)

When the results from a number of sets involving the same number of cooperators are combined, the mean number of hits for the total is the sum of the means of the individual sets, and the variance is the sum of the separate variances.

References

1. Greville, T. N. E. On multiple matching with one variable deck. *Ann. math. Statist.*, 1944, **15**, 432-34.
2. Pratt, J. G. *Towards a Method of Evaluating Mediumistic Material.* Bull. 23, Boston Soc. psych. Res., 1936.
3. Saltmarsh, H. F. Report on the investigation of some sittings with Mrs. Warren Elliott. *Proc. Soc. psych. Res.*, Lond., 1930, **39**, 47-184.
4. Saltmarsh, H. F., and Soal, S. G. A method for estimating the supernormal content of mediumistic communications. *Proc. Soc. psych. Res.*, Lond., 1930, **39**, 266-71.
5. Thomas, J. F. *Beyond Normal Cognition.* Boston: Boston Soc. psych. Res., 1937.

Parapsychology Laboratory
Duke University
Durham, North Carolina

Psychology Department
Duke University
Durham, North Carolina

CURRENT ATTITUDES OF AMERICAN NEUROPSYCHIATRISTS TOWARD PARAPSYCHOLOGY: A SURVEY

By Russell G. MacRobert[1]

Introduction

Ever since the earliest research work in what we now call parapsychology, neurologists and psychiatrists have been prominent in the investigations, and there are indications that they are showing an increasing interest in that field today. Publications on parapsychological topics go back to the times of Richet, Freud, Janet, Stekel, Jung, Mitchell, and McDougall—to mention only a few leading names. There have always been a goodly number of psychiatrists who have contributed to the literature in parapsychology, so that one may well raise the question whether the subject of parapsychology is not of greater interest to psychiatrists than it is to orthodox psychologists.

This possibility is not surprising. The problems of parapsychology have significance for everyone concerned in a professional way with the mind of man, and probably psychiatrists, more than any one else dealing with such matters, are urgently in need of every new finding or angle of approach which will give better understanding of the mystery that still enshrouds many features of man's mental life.

Therefore, to investigate the reaction of psychiatrists to the work going on in parapsychology and to report the results to the parapsychologists as well as to the psychiatrists themselves seemed well worth the effort involved. Accordingly, a questionnaire was directed, between April 15 and August 30, 1948, to 2,510 doctors, specialists in psychiatry and neurology throughout the country, in an effort to find out how they felt about present-day parapsychology. The specialists queried were those listed in the current directory of the American Board of Psychiatry and Neurology dated June, 1947,

[1] Dr. MacRobert is a practicing neuropsychiatrist in New York City. He is a diplomate of the American Board of Psychiatry and Neurology and a member of the Association for Research in Nervous and Mental Diseases.—Ed.

and in the membership roll of the Association for Research in Nervous and Mental Disease, published in April, 1948.

The letter sent out with the questionnaire introduced the topic of the survey in the following words:

"Men of eminent scholarship, unquestioned sincerity, and well trained in the methods of scientific inquiry have been led by experiment to the conclusion that awareness of events, thoughts, or objects can occur otherwise than by any known sensory process or by rational inference and in a degree not ascribable to chance; and that there exists in man an extrasensory capacity which appears not to fall within the scope of what is at present recognized law."

The questions asked were the following:

1. Are you familiar with the recent findings of psychical research pertaining to extrasensory perception? (Yes—No—Slightly)
2. Do you believe that personal bias, experimental or statistical errors may account for the positive findings? (Yes—No—Uncertain)
3. Do you believe that psychiatrists and neurologists would serve a useful purpose by sponsoring research to determine if extrasensory perception has a place in the psychodynamics of the nervous system? (Yes—No—Undecided)
4. Have you ever observed, in your general experience or professional practice, anything which would indicate an extrasensory awareness? (Yes—No—Uncertain)
5. Do you feel that identification with psychical research would discredit your professional standing? (Yes—No—Uncertain)

Results

Doctors share the popular aversion to questionnaires, but 723 of them, or 28.8 per cent, had returned answers by October 18. Among the respondents there appeared to be most of the active, eminent members of the specialty.

The general results, showing how the answers were divided on each question, are given in Table 1 in terms of both the actual numbers and as percentages.

It is interesting to inquire how those who replied a certain way on any one of the questions answered the other four. For example, how do those who say they have observed something indicating an extrasensory awareness reply to the other questions? The figures

Table 1

TOTAL RESULTS—723 REPLIES

Questions	Yes	No	Indeterminate
1. Familiar with findings?	222 (31%)	86 (12%)	415 (57%)
2. Bias or errors the explanation?	274 (38%)	125 (17%)	324 (45%)
3. Useful purpose to sponsor research?	495 (68%)	128 (18%)	99 (14%)
4. Personally observed ESP occurrences?	163 (23%)	437 (60%)	123 (17%)
5. Identification discredit standing?	138 (19%)	453 (63%)	132 (18%)

for this special group are shown in Table 2. Personal experience with ESP does appear to be related to the other opinions expressed. This is a general finding that is suggested by comparing the percentages of those who have observed ESP occurrences (Table 2) with the results for the entire group (Table 1). We find that a larger percentage of those who claim direct experience with parapsychological occurrences say that they are familiar with the findings (Question 1) and that their professional group should sponsor parapsychological research (Question 3), but a smaller proportion of them say that bias or errors can explain the positive findings (Question 2) and that identification with psychical research would discredit their professional standing (Question 4).

Table 2

INDICATES HOW 163 WHO ANSWER "YES" TO QUESTION 4 (THAT IS, THEY CLAIM EXPERIENCES INDICATING AN EXTRASENSORY AWARENESS) ANSWER THE OTHER FOUR QUESTIONS

Questions	Yes	No	Indeterminate
1. Familiar with findings? 64 (39%)		16 (10%)	83 (51%)
2. Bias or errors the explanation?	22 (14%)	69 (42%)	72 (44%)
3. Useful purpose to sponsor research?	151 (92%)	6 (4%)	6 (4%)
4. Personally observed ESP occurrences?	163	.....	.....
5. Identification discredit standing?	25 (15%)	103 (63%)	35 (22%)

Although these comparisons in terms of percentages are interesting, the information they give is not as definite as one would like to have. They suggest certain relationships, but they do not provide an adequate basis for judgment. Before we can reach any conclusions on the strength of these observations or others like them, it is necessary to apply some sort of statistical evaluation to see whether the differences indicate more than merely chance fluctuations. Some method is required to show whether the answers on the different questions are correlated with each other in any significant way.

When an appropriate statistical test was applied, several instances were found in which the manner of replying to one question was significantly correlated with the way of responding to another. Among the significant relationships found were the following:

Those who claim experiences indicating extrasensory awareness are more likely than the others to say that they are familiar with the findings.

Also, those who have observed ESP are more likely to believe that the positive findings cannot be attributed to bias or errors.

Again, those claiming experiences tend to believe that psychiatrists and neurologists would serve a useful purpose in sponsoring research in parapsychology.

Those who are open-minded toward the past work (the ones answering "no" or "undecided" when asked if they believe that bias or errors can explain the recent findings) tend to believe that psychiatrists and neurologists can serve a useful purpose in sponsoring research.

Those who believe that bias or errors can explain the positive findings are more likely to believe that identification with the research would discredit their professional standing.

Finally, those who say that psychiatrists and neurologists would serve a useful purpose in sponsoring research tend not to believe that identification with parapsychological research would discredit their professional standing.

For the benefit of the more statistical-minded readers, it seems worth while to describe more fully the type of analysis used in arriving at the above results. This is presented as an appendix.

The comparison of the answers to the question regarding familiarity with the research (Question 1) and attitude toward the positive findings (Question 2), as shown in the appendix, presents some puzzling aspects. On the one hand, there is a significant positive correlation that seems to suggest that those who claim familiarity with the research are more likely to attribute the findings to bias or error. On the other hand, we see that 172 who said they were not familiar with the findings or were only slightly so also said that they believed bias or error was the explanation. Their answers are logically inconsistent, as anyone who is not familiar with the research is not qualified to pass critical judgment upon it. The fact that so many of them did so anyway raises the question whether many of those who answered "yes" to both questions may not have done so to avoid this inconsistency. That is, holding the a priori opinion that bias or errors can explain the results, did they claim familiarity with the research to justify themselves in rejecting the recent findings? In the absence of any objective evidence of how the different respondents interpreted the word "familiar," this at least can be offered as an interesting speculation. Question 1 was purposely left vague in order to provide a good opening wedge and thus to encourage as many as possible to fill out the questionnaire.

Written Comment

Besides checking their questionnaires, many respondents sent in written comment which altogether amounted to about seventy typewritten pages when deleted or identifying material. Some of the representative comment follows, with personal and circumstantial references deleted in order to preserve anonymity.[2] The numbers in parentheses at the beginning of each quotation indicate how the respondent answered the five questions. The figure "1" indicates "yes"; "2" indicates "no"; and "3" indicates "slightly" or "uncertain." Thus the index number (31221) means that the respondent answered "slightly" to the first question, "yes" to the second; "no" to the third and fourth, and "yes" to the fifth.

Question 1

By its very nature this question elicited very little comment.

[2] On the recommendation of the editorial staff a large body of favorable comment of a general nature was omitted to conserve space.

Question 2

Among those who answered "yes" to Question 2 (Do you believe that personal bias, experimental or statistical errors may account for the positive findings?) were a number who believed that faulty statistics could explain the results. These are represented by such a statement as the following:

> (31221) His [Rhine's] statistics show that his subjects are correct slightly more often than could be expected by sheer chance. It is my opinion that while believers in ESP are sincere, they are misinterpreting their statistics.

The most commonly offered counterhypothesis, however, was that there must be some way of explaining the ESP results on a sensory basis. The following quotation is representative:

> (31321) The unconscious part of the mind—as the result of experience, both of the race and the individual—often gives us "hunches" or premonitions of coming events; but these are delayed sensory responses, or responses derived from sensory impressions which have been perceived but not noted at the time by the conscious mind. These responses are so surprising at times that we get ideas of an extrasensory faculty.

Another went so far as to say:

> (11221) It seems probable that most belief in ESP arises from the well-known symptom of delusion of alienation of patients suffering with schizophrenia.

Various other counterhypotheses were offered, such as coincidence, credulousness, deja vu, and even charlatanism. One said:

> (11122) It is not psychical research. It is physiology.

Whatever their basis for unfavorable criticism, most of these commentators displayed an open mind on the subject, maintaining, in spite of their disbelief, that there is no good to accrue "from a reluctance on the part of scientific men to test fully any notions which appear to be susceptible to the scientific method" and that "the only way we can advance human knowledge is by having an open mind and admitting that anything is possible in the psychic field." This point of view is well expressed in the following comments:

> (32332) I do not believe that the attempt to discount positive findings on grounds of "personal bias" of the investigator,

etc., is intellectually honest. I have no doubt that the findings are statistically significant."

(31132) My own impression of positive results so far in ESP is that unknown and uncontrolled variables have entered into the experiments. Nevertheless, since belief in ESP plays such a great role in religion, history, biography, cults, and everyday social intercourse, it would seem to be important to discover any possible basis for this belief. As Samuel Johnson has observed in regard to literature, anything which retains the public's favor over a long period of time must have something which the experts have overlooked.

From those who did not believe that bias, etc., could explain the results, there was a considerable body of favorable expression, such as the following:

(23112) . . . I am greatly interested in this phase of psychology, for I "feel" that there is a medium of communication other than what is evident and obvious today. There is so much to be known—in fact no person today knows the normal variations in the cytoarchitecture of the human cerebral cortex, much less the comparative anatomy, the patho-physiology, the physiological psychology, . . .

(33222) Most of the investigators in this work seem people of integrity. Personal bias is a poor explanation and should be considered only after other explanations have been exhausted.

And another, more typical one:

(13112) It is indeed pleasing that someone is fathering an investigation of parapsychology in a scientific way. A lot of present medical facts were once hidden behind mystery.

Question 3

As was pointed out before, there were very few of the respondents who did not maintain an open mind on the matter of research into ESP. Question 3 (Do you believe that psychiatrists and neurologists would serve a useful purpose by sponsoring research to determine if extrasensory perception has a place in the psychodynamics of the nervous system?) brought greater unanimity than any other. It was largely agreed that ESP phenomena should be studied and that psychiatrists are in an extremely advantageous position to cooperate and to initiate studies in these fields. Some typically favorable comments follow:

(32112) I have been confronted with far too many occurrences along the line of this questionnaire, both with patients and personally, to be able to dismiss these clinical facts with

a shrug, or by attributing them to chance. I think it in keeping with our scientific heritage to interest ourselves in an attempt to explain such observable phenomena. For myself, the only thing which keeps me from giving it more attention and thought is a combination of lack of time and a feeling that the implications are so vast that I wouldn't know where to begin. I would applaud our profession putting its best minds to investigating along these lines.

(32113) It seems to me that psychiatry is guilty of ignoring a vast and fruitful field for investigation. Our knowledge is not complete until we know more about ESP. The Oriental psychologists are far ahead of us in this field.

(13112) I think it is time for neuropsychiatrists to take some organized and official interest in this ESP work.

(31122) I believe it would be a good idea to sponsor research in parapsychology and let the findings come out as they will. I think this type of investigation should be carried on as free as possible from any preconceived bias. I would therefore favor a research program supporting research on this subject but with no strings attached.

(12112) . . . For a great many years I have thought definitely that there is such a thing as ESP and I have had a number of patients whom I suspected possessed this sense. I believe this problem should receive the support, study, and research of psychiatrists and neurologists.

As a means of implementing such an investigation there were various suggestions:

(32332) I believe that collaboration between laboratory investigators and clinicians in psychiatry and neurology might be planned in such a way as to give valuable information. . . . It might be of value if clinicians could refer . . . patients to laboratories where it could be determined objectively whether or not such patients make consistently high scores as compared with control groups.

(33112) Psychiatrists, and particularly psychoanalysts, could record carefully any experiences they may have in their practice and report them periodically to a committee.

(23113) . . . scientific research and control should be done under the supervision of a board of recognized authorities in different medical specialties.

(32123) It is not necessary to make a financial contribution or to take time to do experiments,—but a symposium at a scientific meeting where the most reliable proponents might state their case without being prejudiced as frauds would be valuable.

Specific suggestions as to research were made by some:

(33132) I would say that I believe so-called coincidence in thought timing between two people occurs so frequently that it

should be investigated. As for the so-called unreliability of mediums, I am willing to consider Freud's comment that it is possible that only persons who are unreliable—that is, whose minds are perhaps less accurate than ordinary people—may have the quality which enables them to register occurrences without sense perception.

(11312) Collecting data on spontaneous events and sifting these out may be all that is possible now.

(31122) I am . . . interested in a research study on *dreams* which of course could indirectly throw light on the problem you are interested in.

(32332) It would . . . seem to be valuable to conduct parallel longitudinal studies of selected patients—that is, regular laboratory studies carried out independent of but parallel to psychodynamic (or analytical) studies of the same person. This type of investigation might eventually yield information concerning the role of emotional factors in determining the degree of extrasensory awareness and possibly might help to differentiate between greatly increased sensitivity to sensory stimuli and genuine extrasensory perceptions.

It would also seem to me of some value to have a standard form for reporting observed experiences prepared and sent to interested clinicians so that future experiences could be observed and recorded and forwarded to a committee for evaluation.

(31122) I would initiate the research by extensive study, thorough analysis, of the more recognized so-called mediums who appear to form the basis of a good deal of the psychic discussions.

(33112) It is hard to figure out a specific research program but it seems to me that the part played by telepathic communication between doctor and patient would be favorable in relation to its unconscious perception on both sides.

(33112) Monozygotic twins think more alike and "can read each other's thoughts" more readily than dizygotic. Monozygotic twins have identical brain waves. Is this related to their thinking alike? Do they score higher than dizygotic twins in the ESP card game? These points deserve testing.

(13122) We should especially study the personalities of the individuals who seem to possess this capacity.

Those who took a dim view of the advisability of sponsoring research did so in some cases because they felt that psychiatrists were not sufficiently interested or could not afford the time or because they thought the investigations should be left to the parapsychologist:

(11222) I think psychiatry and neurology have enough to do in the field of mundane and sensory perception. Let's stick to our job. Let others work in this field.

(31122) I think any disputed point should be investigated—but I think there are so many more important matters to occupy the energies of neurologists and psychiatrists that I could not favor *sponsoring* research in this field—for those who have the urge, the light is green.

(11211) No. Let the parapsychologist continue his work, but let him study the reality of other psychic phenomena which he now ignores.

Question 4

Question 4 (Have you ever observed, in your general experience or professional practice, anything which would indicate an extrasensory awareness?) brought answers varying all the way from the respondents who claimed never to have observed the occurrence of ESP to the following one who claimed to have observed the full range of phenomena:

(11211) Yes. I have seen all types of physical and mental phenomena, veridical and evidential, produced by accredited psychics.

A few of the respondents had observed the occurrence of ESP in experiments of their own devising, such as the following:

(12112) I have been able to convey simple commands (such as "sit down," "stand up," "turn around") simply by willing these commands—to a ***hypnotized*** patient.

In the written comment there was a body of case material of a personal or professional nature from which several selections have been made:

(33111) A five-year-old girl awoke her mother, who was sleeping in the same room, it being the nursemaid's 24 hours off, and said, "I don't want to die and go far away from you, mummie." The mother picked her up and held her a few minutes and assured her she was going to live to grow up and be a "mummie" too, etc. The next morning the mother told me about this and asked me to speak to the colored nursemaid and said, "I'll not keep her if she tells the children tales that frighten them."

After lunch the same day, the children lay down for a nap. The little girl sat up at once, called her mother and said, "Hold me close, I feel funny." The mother took her, sat down and began to rock her, and said, "Where do you feel funny dear? What do you mean?" No reply, so she gently put her to bed, asleep she thought; but the eyes were only half closed and she could find no pulse. The child was dead. Autopsy showed a greatly enlarged heart.

(11312) One of the most startling experiences I know of was by a social worker colleague; returning by rail from a convention early in the morning she woke up with a terrifying nightmare in which she saw a scene of wreckage, dead bodies and ambulances. She could not return to sleep and woke up a friend to sit with her. On the afternoon of the same day, without any advance word brought to them on the train, they passed almost the identical scene, where a train had been derailed the night before.

(13111) I had had lunch in my office with a friend in the publishing business. His name was "C." We had had a conversation of great interest to me. I then went by taxi to Blank Hospital to call on a patient. . . .

When I entered the patient's room he was obviously very glad to see me. He looked even more cachetic and cadaverous than before. Since I realized my own powerlessness to help him, I was perhaps a little more "personal" and "sympathetic" than usual. He said something like this:

"There is only one thing wrong with your visits, Doctor. You don't stay long enough." He then went on: "When Doctor — operated on me," but instead of "H" he inserted the name "C," the very name of my publisher friend whom I had just lunched with. "C" has a rather unusual name. There are only sixteen of them in the New York telephone book. I was astonished to hear my patient pronounce it—so much so that I asked him why he had said "C." He said: "I don't know—it just popped into my head."

From where? This patient knew nothing about me or my associates.

Question 5

Question 5 (Do you feel that identification with psychical research would discredit your professional standing?) brought out both extremes of comment. Those who answered "yes" were in the minority and their convictions seemed usually to have been induced by expediency, as the result of such situations as the following:

(32111) I believe that at the present time, identification with psychical research would undoubtedly discredit my professional standing in varying degrees. . . . However, this is not to imply that I consider the fostering of such a study to be unwarranted; nor that I am unsympathetic to the cause and to its purpose, as I think it to be.

Perhaps it would be well to align a number of prominent neuropsychiatrists on the positive side of such an endeavor so that those of us who bask in the reflected light of our profession could cite the names of such prominent men to justify, if need be, our membership in such a group.

The large majority, however, agreed with the following commentators who thought that their professional standing would not

be jeopardized by connection with well-planned research done under reputable leadership:

> (12123) . . . it would depend entirely on the nature of the research and those supervising and participating in it. I would see no reason why research into psychical phenomena could not be conducted with as high an ethical standard as any other research, if such standards were originally set up and adhered to.
>
> (33122) Why should such research discredit a man's professional standing? I assume that ESP, if it exists, is a function of the human mind, and if so, who is better qualified to investigate it than psychiatrists and their professional assistants? I have no axe to grind for ESP, but it is no more startling than Freud's investigations must have been when first they were begun.
>
> (13132) There appears to be a tacit condemnation of phenomena not yet understood, and condemnation without investigation must keep us in darkness. This was the prevailing attitude of the middle ages—an attitude of mind of which Shakespeare wrote, "There is more in Heaven and earth, Horatio, than is dreamed of in your philosophy. . . ." Honest investigation to my mind is the only answer and should bring no discredit to anyone's standing, professional or otherwise.
>
> (11312) I am not worried about my professional standing, as much as remaining interested in challenging and significant phenomena.

Remarks

It must be realized that no verdict upon the work of the parapsychologist can be considered rendered by this report. No verdict on the merit of the work was sought. Such verdict must be based upon study of the original reports of research. As the survey was an attempt merely to assess current attitude, the word "familiar" in Question 1 was deemed sufficiently vague to admit casual students of the work.

With this in mind, the tables and comment speak for themselves. Especially noticeable is the fact that 495 doctors, or 68 per cent of the respondents answered "Yes" to Question 3, which concerned the sponsoring of research.

The author is grateful to those doctors who assisted the survey by completing the questionnaire and supplying the comment. Despite the fact that this survey was not the usual mere sampling procedure of public opinion polls, but a complete canvass of the

membership roll, and even though the reply percentage is what would be considered high by professional pollsters, the writer is disappointed that all questionnaires were not returned. The certificate of the American Board of Psychiatry and Neurology is analogous in a measure to the qualifying degree in medicine. If a doctor is ill, inactive, or retired, he is not deprived of his listing. How large a proportion of those unheard from fall in this latent category and not merely in the disinterested group, is not known. Also, the directory of the American Board of Psychiatry and Neurology does not yet contain the names of many younger specialists. Their attitude to this new branch of psychology remains unknown.

At present relatively few psychiatrists are familiar with the recent findings of parapsychology, judging from the answers to Question 1. A wider dissemination of the information is extremely important so that at least informed judgment may be passed on the investigations.

Appendix

For convenience in making the statistical analysis used to see whether the attitudes expressed in replying to the five questions were related to one another, the answers to each question were considered as belonging in either of two categories: the "yes" replies were taken as one group; and the "no" and noncommittal replies were combined as a second group.

For each part of this study, it was observed how each respondent marked the two questions to be compared, and a tally mark was placed in a two-by-two table to show his particular combination of answers. For example, the form of the two-by-two table for the answers to the first two questions had this appearance:

		Question 1	
		Yes	No or Slightly
Question 2	Yes		
	No or Uncertain		

Suppose that the first questionnaire examined showed that the respondent answered "yes" to both of these questions. A tally mark was made in the upper left cell of the table. Suppose that the second respondent replied either "no" or "slightly" to the first question and "yes" to the second one. His case fell within the upper right cell. This process of observing and tabulating was continued until all 723 questionnaires had been properly entered in the contingency table.

This kind of tabulation was done for all ten possible ways of comparing the five questions, taken two at a time. The analysis thus led to ten two-by-two contingency tables, each of which showed how the replies to a designated pair of the questions were related to one another. A convenient way of presenting these ten tables is to show them in a single large table, with the questions concerned given along the top and left-hand margins. Under each marginal question appear the two classes, into one of which each of its answers must fall. This large table, when completely filled out, will present the contingency table for every possible comparison of one question with another. To find the figures for any particular pair of questions, it is necessary only to locate these two questions on the two margins and to find the table at the point of their intersection within the larger table. The ten contingency tables for all the question-to-question comparisons are shown in the appendix table.

A further word of explanation is required in regard to the data shown in this table. The figures not in parentheses in each cell are those for the data actually observed. The figures given in parentheses indicate the most probable number of cases expected in each instance if the replies to the two questions have only a random or chance relationship to one another. The method of evaluation most commonly applied to such a two-by-two contingency table is the chi-square test of significance. This measures the degree to which the observed distribution departs from the expected one. The evaluation itself need not be explained here except to say that a chi-square value of 6.63 or higher is customarily taken as indicating statistical significance, or a significant correlation in either the positive or negative direction. If a table yields a significant chi-square value, this indicates that the result would not happen by chance more often

ANALYSES TO SHOW WHETHER ANSWERING EITHER "YES" OR "NO OR NONCOMMITTAL" WAS EITHER POSITIVELY OR NEGATIVELY CORRELATED FROM QUESTION TO QUESTION

Question 1	Question 2		Question 3		Question 4		Question 5	
Familiar with findings?	Bias or error the explanation?		Useful purpose to sponsor research?		Personally observed ESP occurrences?		Identification discredit standing?	
	Yes	No or uncertain	Yes	No or undecided	Yes	No or uncertain	Yes	No or uncertain
Yes	102 (84)	120 (138)	154 (152)	68 (70)	64 (50)	158 (172)	51 (42)	171 (180)
No or slightly	172 (190)	329 (311)	341 (343)	160 (158)	99 (113)	402 (388)	87 (96)	414 (405)
	$X^2=8.82$ Significant positive correlation		$X^2=.12$ No relationship		$X^2=7.24$ Significant positive correlation		$X^2=3.13$ No relationship	
Question 2 Bias or errors the explanation?	Yes		146 (188)	128 (86)	22 (62)	252 (212)	86 (52)	188 (222)
	No or uncertain		349 (307)	100 (142)	141 (101)	308 (348)	52 (86)	397 (363)
			$X^2=47.09$ Highly significant negative correlation		$X^2=53.24$ Highly significant negative correlation		$X^2=43.22$ Highly significant positive correlation	
	Question 3 Useful purpose to sponsor research?	Yes			151 (112)	344 (383)	53 (94)	442 (401)
		No or undecided			12 (51)	216 (177)	85 (44)	143 (184)
					$X^2=56.96$ Highly significant positive correlation		$X^2=71.38$ Highly significant negative correlation	
		Question 4 Personally observed ESP occurrences?	Yes				25 (31)	138 (132)
			No or uncertain				113 (107)	447 (453)
							$X^2=1.92$ No relationship	

than one time in a hundred such comparisons. If the chi-square value is as large as 40, as many of those in the table are, the odds against the chance explanation are millions to one.

It is a simple matter to tell whether the correlation in each case is positive or negative. To find out, it is only necessary to compare the two figures in the upper left-hand cell of the contingency table in question. If the number of observed cases in that cell is *larger* than the number expected, the correlation is *positive.* This means that there was a tendency for each respondent to answer both questions the same way, either "yes" to both or "no or noncommittal" to both. If the number of observed cases in any "yes-yes" cell is *smaller* than the number expected, the correlation is *negative;* there was, in other words, a tendency for each respondent to give replies in opposite categories on the two questions concerned.

555 Park Avenue
New York 21, N. Y.

A NEW METHOD AND AN EXPERIMENT IN PURE TELEPATHY

By WILLIAM R. BIRGE

ABSTRACT: Three series of pure telepathy runs with ESP cards were conducted under conditions which were devised to eliminate the occurrence of clairvoyance. The scores on the direct target were not above the level of chance expectation, though strongly suggestive scores were obtained when each of the subject's calls was checked against the target symbol which came just after that for which it was intended.

The author also describes a new method of checking telepathy card-calling data. This method involves the use of a code thought of by the experimenter and communicated to the checker in an objective manner, but in such a way that there is a minimum possibility of the subject's using clairvoyance to ascertain the target symbols.

Mr. Birge is a graduate assistant in the Psychology Department of Duke University, and he has been a member of the Parapsychology Laboratory since 1946.—Ed.

PART I. THE EXPERIMENT

IN 1945 Rhine (**7**) challenged the status of telepathy as an established phenomenon. On the grounds that conclusive evidence existed for both clairvoyance and precognition, Rhine argued that the existing evidence for telepathy could be explained in terms of one or both of these capacities. He concluded from this state of affairs that telepathy could be demonstrated only by rigidly excluding all possibilities of clairvoyance. The target symbols in a pure telepathy test must exist only in the mind of the agent, and they must not at any time be objectively represented in either speech or writing.

At least three parapsychologists, McMahan (**5**), Soal (**9**), and Thouless (**14**), have conducted experiments in pure telepathy which satisfy the conditions as set forth by Rhine. McMahan and Soal have reported positive results indicating the occurrence of telepathy in their investigations. Thouless has not as yet published the results of his experiment. The present report is based on the fourth independent research in pure telepathy which has been conducted since the conditions for its demonstration have been so radically revised. The results of this study provide suggestive confirmatory evidence of telepathy.

General Description of the Procedure

The experiment consisted of three series of pure telepathy tests. A series was composed of 12 experimental sessions in each of which a subject attempted to call the order of four decks of ESP cards through telepathic communication with the agent (W.R.B.). In this way, 48 runs were contributed during each series, while the three series combined yielded a total of 144 runs of pure telepathy data.

Before the formal testing began, the experimenter mentally devised a code whereby each of the five ESP symbols was associated with a number from one to five. The purpose of this code was threefold: (1) to permit a random selection of target symbols to be made; (2) to simplify the sending and checking processes by allowing the agent to make use of cards and record sheets; (3) to allow an independent recheck of the data. Naturally the code could not be revealed at any time if the "pure" character of the results were to be maintained.

Four decks, each containing 25 code cards, were utilized in the experiment. These decks were prepared by stamping the faces of blank cards with the numbers one through five. Each deck contained five of each of the code numbers.

For Series I and III, the order of the target symbols for all 12 sessions was determined shortly before the commencement of each series. The target order for an individual session was chosen by shuffling the four decks of code cards and recording their order in the first four "card" columns of an officially numbered record sheet and its duplicate. When 12 sets of record sheets had been prepared in this way, they were given to an assistant who shuffled them and renumbered each pair of record sheets one through twelve. The originals were returned to the agent and the carbon copies were deposited with the laboratory librarian, Dr. Humphrey. The precaution of depositing one set of record sheets with the librarian was taken in order to make it impossible for the agent to alter the order of target presentations during the series.

The record sheet marked #1 was used with the first subject of the series, the sheet marked #2 with the second subject, and so on. Before each experimental session, the agent arranged the four decks

of code cards to match the order recorded on the proper record sheet. For each trial during the session, he exposed the top card of the target deck, thought of the associated ESP symbol, and then placed the card face down beside the code deck.

As mentioned before, an assistant shuffled and numbered 12 sets of record sheets for Series I and III. For Series II, however, the assistant received four duplicate sets of record sheets to shuffle and number before the first session, and eight duplicate sets of record sheets to shuffle and number just prior to the fifth session. The assistant was given the task of shuffling the record sheets and thus determining the session order in which they were to be used in order to eliminate the possibility of a precognitive ESP shuffle by the agent. If the latter had shuffled and numbered the target record sheets, he might have arranged them in such a way that extrachance results of a nontelepathic character would occur through precognitive clairvoyance of the subjects' calls combined with an ESP shuffle. The ESP shuffle hypothesis could not apply to the shuffling of the record sheets by the assistant since the latter was ignorant of the agent's code.

Agent and subject were separated by not fewer than two rooms during every experimental session. The agent notified the subject that he was ready to send a symbol by pressing a key which turned on a small lamp both in his and in the subject's room. After the subject had recorded his call, he pressed a similar key which turned out the lights in both rooms. The signaling system permitted no sensory cues of the target symbols to be transmitted, since the lamps in both rooms could be turned on only by the agent and turned off only by the subject.

The subject recorded his calls in the first four "call" columns of an officially numbered record sheet with a carbon duplicate. At the end of the session, the agent would take the subject's original call sheet and make a silent check of the total number of direct hits scored on the four runs. The subject was then asked to record this total hit score in the appropriate place on his carbon call sheet. Still in possession of the carbon call sheet, the subject was then accompanied to the room of a laboratory assistant who deposited the carbon record sheet in a locked box. The agent retained the original call sheet for his own files.

All but one of the subjects were men and women undergraduate students at the university. In Series I and II the subjects were unselected, with the exception of one subject whose first score was so high that she was asked to participate in seven additional sessions. The subjects for Series III were selected on the basis of their excellent performances in a classroom test under the supervision of Dr. Burke M. Smith.

While the intricacies of the code were not explained, the subjects were all informed that they would be tested for telepathy. Each subject was told that when his light came on he should guess the ESP symbol which was being thought of by the agent.

For the first six sessions of Series I the agent entered the subject's room after each run to inquire if the run had been completed satisfactorily and to offer a word of encouragement. The main purpose of this break in the testing period was to minimize fatigue and monotony. The agent did not look at the subject's record sheet during these interruptions other than to make sure that the correct number of trials had been recorded in the proper places.

Beginning with the seventh session of Series I and continuing through the remaining sessions of the entire experiment, the intervals between each card-calling run were taken up with a short PK dice test introduced for diversification. These tests gave insignificant results.[1]

One of the standard precautions in parapsychological research is the independent rechecking of all experimental data. This is difficult to accomplish in pure telepathy research since the agent's code must be communicated to another person for the recheck to be made. If the code is transmitted directly, it is theoretically accessible to the clairvoyance of the subjects.

[1] At each trial in these tests the subject threw six dice from a cup onto a table with padded surface and sides. In Series I neither the number of PK runs nor the order of target faces was controlled in a systematic manner. In Series II and III, each subject threw 12 runs for the same target face. Four runs were thrown during each of the three intervals between card-calling runs, and the order of target faces was rotated so that the first subject threw for ones, the second subject for twos, and so on through the sixth subject, after whom the cycle was repeated.

The direct deviation score for 288 PK runs was —5. When the hits were analyzed in terms of the four quarters of the page, the following figures were obtained: 1st quarter, +3; 2nd quarter, —21; 3rd quarter, +8; 4th quarter, +5. The deviation scores are each based on 72 runs, and neither individually nor comparatively do they fall outside the range of chance expectation.

During the course of this experiment I devised in collaboration with my colleagues a new method for the rechecking of pure telepathy data. Because of its complicated nature, the method will not be discussed in detail until a later section of the report. It is sufficient to mention here that the data were independently rechecked, while the strictest precautions against clairvoyant leakage were maintained. In this rechecking I had the assistance of Miss Elizabeth McMahan and Dr. Betty M. Humphrey.

Results

The deviation scores on the target cards were as follows:

Series	Runs	Dev.
I	48	− 7
II	48	+12
III	48	−22
Total	144	−17

These results fall clearly within the range of chance expectation. It is evident that this analysis of the direct scores offers no evidence of the operation of telepathy.

A chi-square significance test was applied to the data, in which the direct scores of each subject on four runs were converted into CR's. The CR's were then squared and summed, yielding a χ^2 for the three series combined of 35.41. With 36 degrees of freedom, this χ^2 gives a probability value of .496. The distribution of direct scores is thus in accordance with chance expectation.

The distribution of direct hits in terms of the page quarters also revealed nothing of significance. In this analysis the middle segments of 20 trials on each page were arbitrarily omitted in order to obtain an equal number of trials in each quarter. For the three series combined the following figures were obtained:

Quarter	Runs	Dev.
1st	28.8	− 1
2nd	28.8	+ 8
3rd	28.8	+ 2
4th	28.8	−17

A chi-square analysis of the quarter deviations yielded a χ^2 of 2.4, with three degrees of freedom. This is clearly insignificant.

It was in the analysis of the displacement scores that the sug-

gestive evidence for telepathy was found. In this study the subjects' calls were matched with the target symbols which appeared just previous to and just after the direct target symbol. For example, a subject was credited with a postcognitive hit if, in the target sequence of circle-cross-star, he called "circle" at the time when "cross" was the actual target. The same subject would have obtained a precognitive hit if he had called "star" when "cross" was the target. The results of the displacement analysis are given in Table 1.

Table 1

DISPLACEMENT SCORES FOR EACH SERIES

	SERIES I		SERIES II		SERIES III		TOTAL			
	Runs	Dev.	Runs	Dev.	Runs	Dev.	Runs	Dev.	CR	CR_d
Precog.....	46.08	+18.60	46.08	+19.60	46.08	+31.60	138.24	+69.90	2.97	2.92*
Postcog....	46.08	−18.40	46.08	− 7.40	46.08	− 1.40	138.24	−27.20	1.16	

*This CR_d has a tabulated probability value of .0018. It is not independent of the CR for the precognition scores, and thus lends only a slight additional weight to the suggestive evidence for telepathy provided by the precognitive displacement scores.

The CR of 2.97 for the precognitive displacement scores has a tabulated probability value of .0015. Since the analysis of displacement scores had not been planned, there is no logical basis for arriving at a correction to be applied.

Discussion

The direct scoring on the target symbols revealed no evidence of the occurrence of telepathy in the present experiment. When the subjects' calls were checked for displacement, however, it was found that the precognitive displacement scores and the difference between the precognitive and postcognitive displacement scores strongly indicate that the results were due to factors other than chance.

The observed displacement scoring cannot be explained in terms of other "normal" counterhypotheses, such as sensory cues, recording errors, optional stopping, and selection of data. Subject and agent were separated by at least two rooms during each experimental session. The signaling system by which they communicated

with each other permitted no sensory cues of the target symbols, since the system consisted of two lamps which could be turned on only by the agent and turned off only by the subject. The data were recorded on officially numbered record sheets, they were independently checked, and they all have been included in the results described. The length of each series was determined in advance.

Research in parapsychology would be immeasurably advanced if the causes of displacement scoring were known. Through the control and correction of conditions leading to displacement scoring it should be possible to bring the goal of consistent positive scoring into sight. Because of the great importance of this problem, it is worth while to review the literature for possible suggestions concerning the dynamics of displacement scoring.

In addition to this report there is a great deal of evidence indicating that significant displacement scoring is associated with avoidance of the target card by the subject. In 1938 an anonymous physical scientist reported displacement effects which occurred in connection with low target scoring (**15**). Later, Carington (**3**), Soal (**10**), Soal and Goldney (**11**), and Stuart (**12, 13**) conducted experiments in which low target scoring was obtained in conjunction with significant displacement scoring. Bindrim (**1**) found a significant displacement trend in the records of a group of Schmeidler's low-scoring subjects. It should be noted also that Russell (**8**) was able to find no evidence of displacement when he examined the records of eight outstanding ESP experiments.

It seems reasonable to assume that an unfavorable test situation is responsible for the inaccurate operation of psi in the experiments showing displacement. The unfavorable feature may have resided in the subtle interpersonal relations of subject and experimenter, or the test procedure may have aroused some conscious or unconscious frustration in the subject. The literature indicates that an element of frustration existed in many of the displacement experiments. For example, the anonymous physical scientist asserts that his displacement scoring occurred on occasions when he felt ill or fatigued. In the Carington drawings experiment the conditions of great distance combined with a long delayed check may have been sufficiently disconcerting to cause the displacement of the subject's psi aim. The

delayed-check conditions were also present in the Stuart experiment which showed displacement scoring. (It is worth mentioning here that both Rhine [6] and Hutchinson [4] have shown experimentally the depressing effect on target scoring of delayed-check conditions.) Finally we know that Schmeidler's subjects, whose significant displacement trend was found by Bindrim, were asked to identify target symbols which were recorded on a sheet of paper. This may well have seemed to be a far more difficult and hence frustrating task for the subjects than the task of calling target cards arranged in decks.

In the experiment which has just been reported it is difficult to account for the displacement scoring in terms of the frustrating nature of the objective test conditions. It seems probable, however, that the subject-experimenter relations were not conducive to high scoring on the target card. The experiment was conducted in the spring and summer of 1946 when the telepathy controversy initiated by Rhine was at its height. The experimenter was eager to find evidence of pure telepathy, and this very eagerness may have created a state of tension within the subjects which caused the observed displacement of psi aim. Although it is indeed speculative, this hypothesis is not without foundation. In view of the great differences which have been found among parapsychological experimenters with respect to the ESP performances of their subjects, it is reasonable to suppose that the dynamics of displacement scoring are to some extent to be found in the interpersonal relationship of subject and experimenter.

Part II. A Proposed Method for Rechecking Pure Telepathy Data

It should be stated at the outset that the complete method as presented here was not applied to the data provided in my experiment. This was unfortunately impossible to accomplish for reasons which will be stated later. Despite this deficiency the method will first be discussed as it should be applied in card-calling studies of pure telepathy.

By way of review it will be recalled that in order to make his data amenable to rechecking, the agent in a pure telepathy test must communicate his code to some other person. Any objective repre-

sentation of the code clearly destroys the "pure" character of the results by introducing the possibility of some form of clairvoyance. We are thus confronted with the problem of how we can permit independent rechecking of data which satisfy the conditions for the demonstration of pure telepathy.

McMahan used a method suggested by Rhine to solve the problem in the case of her own pure telepathy research (5). She communicated her code to an assistant by referring to common memories which they both shared. According to this method, should the real symbol, "star," be associated in the agent's mind with the code symbol, "five," the association would be communicated by a reference to a common memory which would suggest "five" to the assistant. The expression "cocoanut cream pie," for instance, might refer to such a common memory. The assistant would then insure the correctness of the inferred number by replying with a different memory reference which would signify "five" to the agent.

Soal (9) has conducted an experiment in pure telepathy in which the method of code communication resembles that of McMahan. In conveying his code to an assistant, Soal made use of veiled hints in referring both to real and code symbols. The names of mutual acquaintances constituted the code symbols in Soal's code, while five different species of animals comprised the real symbols. To repeat the fictitious examples given by Soal, a Mr. Rose might be called Mr. Flower, and a certain animal might be signified by saying, "This pelican delighted in water sports." The order of hints for both animals and men were randomized in a written communication to an assistant.

With regard to the methods of both Soal and McMahan, the objection has been raised that they are essentially subjective methods and are consequently of dubious reliability. It is one matter to trust the integrity of an experimenter to the extent that we are confident he will not willfully betray the code in a direct manner. It is quite another matter to trust his judgment in communicating the code in such a way that it is not exposed to the possible clairvoyant perception of the subject. Riddles differ immensely in the ease with which they may be understood. Can we have any assurance that only difficult riddles will be used in pure telepathy research? Certainly a more objective method would be preferred.

The method I am proposing requires that the smallest block of data for which the agent may record a hit total is the page, and that the page shall contain not less than four runs. The reason why the individual run hit totals should not be recorded is that such records might conceivably allow the subject to break the code clairvoyantly. Breaking the code on the basis of a four-run hit total would, at least theoretically, constitute a far more difficult task.

The major step in the proposed method is that an assistant should compile the data in the form of a contingency table after the completion of the experiment. For illustrative purposes, let us assume that the agent's real symbols are the standard ESP symbols and that his code symbols are the numbers one through five. The assistant's contingency table will then reveal how many times the subject responded with each of the five real symbols to each of the five code target symbols. For example, when the code target was "one," the table might show 16 responses of "waves," 14 responses of "circle," 18 responses of "star," 11 responses of "cross," and 12 responses of "square."

With a knowledge of the total deviation score made by all of the subjects, and with the aid of the contingency table, the assistant is in a position to break the agent's code. There will probably be not more than two or three combinations of code and real symbol pairs which will yield the same deviation score as that computed by the agent. Once such a combination has been determined, the assistant must make sure he has extracted the correct code by rechecking at least two pages of data. If the hit totals for these pages coincide with those computed by the agent, the assistant may feel assured that he has successfully broken the code. Of course the assistant must determine the right code mentally—he cannot record it or mention it in words.

A sample contingency table illustrating the method of code communication is given below:

	∧	○	=	∟	+
1	2	0	1	3	0
2	1	1	1	0	1
3	1	0	0	2	0
4	2	1	2	1	1
5	0	1	2	1	1

The table includes the data from a single hypothetical card-calling run. For purposes of making the method clear, let us assume that an assistant checker has been informed by an experimenter that the deviation score made by the subject on this run was +1. The assistant, by means of the contingency table, can now extract several trial codes which will yield a deviation score of +1. He might, for example, extract the following code: 1 = star, 2 = circle, 3 = cross, 4 = waves, 5 = square. In order to be sure that he has extracted the same code as that used by the agent, the assistant must now recheck several runs made by the subject. The correct trial code will be the one which credits the subject with the same number of hits on these runs as that computed by the agent. (In the case of an actual experiment, of course, the contingency table would include a great number of runs, and the elimination of trial codes by the assistant checker would involve rechecking several pages of data, each page including at least four runs, rather than a few individual runs.)

The important principle to remember in connection with this method is that *the agent's code is broken on the basis of all of the data contributed by all of the subjects.* A subject cannot improve his scoring by other than telepathic means, even assuming that he has precognitive clairvoyance both of the total deviation score and of the contingency table. For example, let us assume that the assistant, bearing in mind the total deviation score, looks over his contingency table and concludes that the number "one" must have been the agent's code symbol for "star." Let us assume further that a

subject can make the same discovery by means of precognitive clairvoyance and computational reasoning. Now the subject can only arrive at this discovery by seeing that he will have responded to the code symbol "one" with, say, 66 "stars" in the course of the experiment. He cannot use this knowledge to increase the number of his "star" responses to the "one" target symbol since the number of "star" responses to the "one" code symbol must be considered fixed in order that he may be enabled to break the code. The subject clearly cannot improve his scoring by clairvoyant means unless he has precognition of all possible futures.

In tabulating the subjects' calls in the various cells of the contingency table, the assistant should first enter the calls made on the first run of every page. He should next tabulate the calls for the second run of every page, and so on through the fourth or last run of every page. By constructing the contingency table on the basis of run sequences rather than page sequences, the danger of exposing the code to clairvoyant perception is reduced to a minimum. Were the data tabulated by pages rather than runs, clairvoyant awareness of both the page hit totals and the contingency table might lead to a solution of the code. Thus a contingency table comprised of the data from the four runs of the first page, in conjunction with the recorded hit total for this page, would constitute a simpler means to a clairvoyant solution of the code than would a contingency table including the data from the first runs of all of the pages since there is no recorded hit total for these runs. In our ignorance of the limits of the clairvoyant capacity, it is safer to err on the side of caution.

Once the contingency table has been constructed, the agent should mentally determine the correct deviation score obtained by the subject on the basis of the table. The assistant should not undertake to break the code until he has been informed of this score. A great waste of time and effort might result if the assistant should attempt to break the code on the basis of the deviation score computed by the agent from the raw data. Were the assistant unable to extract a deviation score which coincided with that computed by the agent, the resolution of the discrepancy would probably involve the construction of a new contingency table by the assistant and the rechecking of all of the data by the agent.

The objection may be raised that the proposed method does not prevent the agent from waiting until the experiment is over and then deciding upon a code which will yield the most successful results. The agent cannot make this fraudulent code selection if he is required to note the number of hits (or deviation score) made at each experimental session and file this number with another person. This would seem to be a safe procedure if the number of runs in each session is not less than four.

More serious than the above objection is the criticism that the method does not, after all, completely exclude the operation of clairvoyance. As long as the hit totals for each page are recorded, what is to prevent a subject from breaking the code on the basis of computational reasoning combined with clairvoyant perception of the calls on previous pages or precognitive clairvoyant perception of the calls on subsequent pages? Since the limits of clairvoyance are not known, it is impossible to answer this question in a categorical manner. One can say at most that, without the use of a contingency table, it is extremely difficult to break the code even when the page of calls is in full view. Few people would disagree with Broad's statement (2) that an individual should not be assumed to have powers in an ESP test (beyond ESP) which he does not have on ordinary occasions. In any case the page hit totals must be recorded in order to meet the objection considered above. The agent might record these totals in terms of his own code, but here the subjective factor of the agent's judgment is introduced. There is no assurance that the agent would record the page hit totals in a manner which would not make it even easier for the subject to break the code on the basis of clairvoyance.

The method may appear too laborious to some investigators, but this criticism would apply only to an experiment comprising a large number of sessions. Certainly the advantages to be gained by the method are sufficient to justify the time spent in constructing a contingency table. The method permits the independent rechecking of pure telepathy card-calling data with a minimal exposure of target material to clairvoyant perception. The agent's code is communicated to the checker in a wholly objective manner, since the medium of the code transmission is the data contributed by the subjects themselves.

It was mentioned earlier that the complete method as presented here was not applied to the data of my experiment. The reason why this was not done may now be given. A contingency table including all of the data for Series I had been completed before it was considered advisable to tabulate the calls by runs rather than by pages. It seemed useless to complete the contingency table for all three series at this time since the maximum exposure to clairvoyance had already been made. The subject's calls on the first page (four runs) had been compiled in a contingency table. Arranged in this form, the calls in conjunction with the recorded page hit total constituted the most economical means of solution of the code. To continue further with the contingency table would have been as ineffective as locking the barn door after the horse had escaped.

In view of these considerations it was decided that an assistant should break the code through the use of the contingency table constructed from the data of Series I. I accordingly determined from the table the deviation score that would be obtained on the basis of my code and informed the assistant of this score. The code was successfully broken and the data rechecked. The determination of the correct deviation score by me and the extraction of deviation scores from trial codes by the assistant were, of course, accomplished mentally. Of some interest here is the fact that the assistant found three codes which gave the same deviation score as that computed by me. The selection of the correct code could only be made by the assistant's rechecking the first few pages of the data to determine which of the three possible codes yielded the same page hit totals as those recorded by the agent.

Conclusion

A new method for rechecking pure telepathy card-calling data has been proposed in this report. The chief merits of this method are that it permits the agent's code to be communicated to a checker in an objective manner without revealing the code to the present or precognitive clairvoyant perception of the subject. The method calls for the transfer of the code to the assistant checker on the basis of a contingency table including all of the data contributed by all of the subjects. In this experiment, however, the code was communicated to an assistant through the use of a contingency table which

included only the data from Series I. The assistant's tabulation of the subjects' calls in terms of pages rather than runs was the cause of this failure to complete the contingency table. It is recognized that had the assistant tabulated the calls made on the first run of every page, then the calls made on the second run of every page, etc., the precautions against clairvoyance would have been somewhat more thorough than they actually were in this experiment.

The methodological improvement mentioned above was unfortunately not arrived at until it was too late for it to be applied in this investigation. There consequently exists a remote possibility that the results can at least partially be explained in terms of clairvoyance. However, while there exists some experimental evidence of telepathy, there is a total absence of evidence indicating that clairvoyance can operate to the extent and accuracy with which it must have operated in order to have been effective in determining the results reported here. Of the two hypotheses of clairvoyance and telepathy, the latter appears much more likely to have been the process which accounted for the results. Likewise, it seems somewhat more reasonable to attribute the results to telepathy than to the "normal" counterhypotheses which have already been reviewed.

References

1. Bindrim, E. A new displacement effect in ESP. *J. Parapsychol.*, 1947, **11**, 208-20.
2. Broad, C. D. Comments on Dr. Rhine's "Telepathy and Clairvoyance Reconsidered." *Proc. Soc. psych. Res.*, London, 1946, **48**, 20-25.
3. Carington, W. Experiments with paranormal cognition of drawings. *J. Parapsychol.*, 1940, **4**, 1-129.
4. Hutchinson, L. Variations of time intervals in pre-shuffle card-calling tests. *J. Parapsychol.*, 1940, **4**, 249-70.
5. McMahan, E. An experiment in pure telepathy. *J. Parapsychol.*, 1946, **10**, 224-42.
6. Rhine, J. B. ESP tests with enclosed cards. *J. Parapsychol.*, 1938, **2**, 199-216.
7. ———. Telepathy and clairvoyance reconsidered. *J. Parapsychol.*, 1945, **9**, 176-93.
8. Russell, W. Examination of ESP records for displacement effects. *J. Parapsychol.*, 1943, **2**, 104-17.

9. SOAL, S. G. *The Experimental Situation in Psychical Research.* Ninth F. W. H. Myers Memorial Lecture. London: Soc. psych. Res., 1947.
10. ———. Fresh light on card guessing—some new effects. *Proc. Soc. psych. Res.,* London, 1940, **46,** 152-98.
11. SOAL, S. G., AND GOLDNEY, K. M. Experiments in precognitive telepathy. *Proc. Soc. psych. Res.,* London, 1943, **47,** 21-149.
12. STUART, C. E. An ESP experiment with enclosed drawings. *J. Parapsychol.,* 1945, **9,** 278-95.
13. ———. A second classroom ESP experiment with the free response method. *J. Parapsychol.,* 1947, **11,** 14-25.
14. THOULESS, R. Comments on Dr. Rhine's "Telepathy and Clairvoyance Reconsidered." *Proc. Soc. psych. Res.,* London, 1946, **48,** 15-17.
15. ANONYMOUS. A scientist tests his own ESP ability. *J. Parapsychol.,* 1938, **2,** 65-70.

Department of Psychology
Duke University
Durham, North Carolina

MINOR ARTICLES

AN EXPLORATORY PK EXPERIMENT BASED UPON THROWING TWENTY-FOUR DICE

By EDMOND P. GIBSON

ABSTRACT: Two series of dice-throwing tests were conducted, 24 dice being thrown at a time, for the purpose of testing the occurrence of PK. The total scores were not above chance expectation. When the data were examined for trends found in previous PK data, two effects were noted. One was a U-curve formed by an excess of hits in the beginning and ending trials of the column on the score page in Series I.. The other was a decline in scoring between the first and second halves of Series I.—Ed.

BEGINNING in August, 1946, the writer conducted an exploratory series of PK experiments for the purpose of locating subjects for further individual testing. The procedure consisted of throwing for a specified target face 24 dice (one run) at a time from a conventional dice cup or dictaphone box. The dice rolled down a small chute and came to rest on a confined level surface. The number of hits was scored by the experimenter and subject independently, and the results were entered by the experimenter on the score sheet. This experimental situation has been described in greater detail in an earlier paper (2).

The experiments did not yield significant results as far as the total deviation was concerned. Inasmuch as this was the specific objective of the tests at the start, the data cannot be offered as independent proof of the occurrence of PK. Nevertheless additional exploratory analyses revealed certain findings from this study quite similar to those of earlier reports which did reach the conclusion that PK had been demonstrated. These particular findings therefore suggest that PK may have been operating in this experiment as well, and the data seem worth reporting for the light they may throw upon the problem of how this unusual ability functions.

Procedure

The general procedure was as follows: Each subject made 12 successive throws for each of the six target faces at each session, and the results of the throws for each face were recorded in a column. The session thus comprised a complete round consisting of six columns across the page, a total of 72 throws of the 24 dice. The number of sessions for each subject varied. In this manner, two experimental series were carried through. One of these, Series I, was started with all the subjects throwing for the target faces in the ascending order. After 1,008 runs (14 sessions), it was decided to reverse the order of throwing for the target faces and to continue the series through 14 more sessions. In this way, the possible effect of dice bias on the distribution of success between the left and right halves of the page in the series as a whole was controlled.

Series II was planned after the completion of Series I. Its length was set in advance at 28 sessions, 14 using the target faces in ascending order, and 14 in descending order. In this series, however, each subject was left as free as possible to choose which order of targets he would use in each session.

Results

General Results

The 2,016 runs of Series I gave a deviation of +38. The SD is 81.93 and the CR is .46, which is statistically insignificant.

The 2,016 runs of Series II gave a negative deviation of −84. This result is also insignificant, as the CR is −1.03.

These general results of the two series, together with the total for both series combined, are shown in Table 1.

Table 1

General Results

Data	Runs	Deviation	SD	CR
Series I.....	2016	+38	81.93	.46
Series II.....	2016	−84	81.93	−1.03
Total.....	4032	−46	115.87	− .40

Vertical Distribution

All the columns in each experimental series were pooled to show the vertical distribution of success. Each group of three successive trials within the columns was then further pooled to show this distribution in terms of the four quarters of the column. These results are given in Table 2 for each series and for both series combined. Each quarter of the pooled column for the individual series represents a total of 504 runs.

Table 2

VERTICAL DISTRIBUTION OF HITS IN THE COLUMN

Data	Series I	Series II	Total
First Quarter	+59	−34	+25
Second Quarter	−71	−25	−96
Third Quarter	−23	+ 4	−19
Fourth Quarter	+73	−29	+44
CR_d between (1+4) and (2+3)	2.76	−.49	1.59

It may be seen from the results in Table 2 that the vertical distribution of successes in Series I gave a distinct U-curve, the scores of the first and fourth quarters of the column both being above chance and those of the second and third quarters below. The difference between the first and fourth quarters combined and the second and third combined is 226 hits. The standard deviation of this difference is 81.93 and the CR, 2.76. The chance expectancy that there will be a U-curve as marked as this is, therefore, .003.

Series II, on the other hand, did not show the same vertical distribution. It is interesting to note, however, that the distribution observed is in the direction of giving an inverted U-curve, the two end quarters of the column producing a total negative deviation of 63 hits and the two middle quarters a negative deviation of 21 hits. The difference between the end and the middle quarters in Series II is, of course, not significant.

When the vertical distributions of Series I and Series II are combined, they tend to cancel one another out, though the total is still slightly in the direction of a U-curve.

Horizontal Distribution

A number of previous studies have shown that there is a general tendency for the PK scoring rate to decline as the subject proceeds across the record page (8). The present results were analyzed to show the average rate of scoring in each of the six column positions, from left to right across the page, to see whether a similar decline occurred in the present work. The results, shown in Table 3, indicate that no such decline in scoring occurred.

Table 3

AVERAGE RUN SCORES FOR THE SIX COLUMN POSITIONS ACROSS THE PAGE

Data	Columns					
	First	Second	Third	Fourth	Fifth	Sixth
Series I.....	4.04	4.08	4.01	3.93	4.07	3.98
Series II.....	4.15	3.82	3.85	3.89	4.01	4.04
Total.....	4.10	3.95	3.93	3.91	4.04	4.01

Chronological Decline

A further analysis was made to determine whether there was any consistent chronological trend in scoring as the subjects proceeded through the period covered by the two experimental series. First, each individual series was divided at the midpoint in terms of the number of runs involved, and the difference between the first and second halves was computed. Finally, the difference between Series I and Series II was taken as reflecting a possible chronological trend extending over the entire experimental period. The results shown in Table 4 indicate that a decline in scoring did occur within Series I. The 1,008 runs of the first half of Series I, in which the ascending order of targets was used, gave a deviation of +110. The second half of the series, in which the target faces were thrown in descending order, gave a deviation of −72. Thus, the chronological decline between the two halves gave a difference of 182 hits. The SD of the difference is 81.93, and the CR is 2.22. A drop in scoring as great as this would be expected 13 times in 1,000 such series on the basis of random sampling.

Table 4

CHRONOLOGICAL DISTRIBUTION OF HITS

Data	First Half	Second Half	CR_d
Series I: First Half vs. Second Half	1008: +110	1008: −72	2.22
Series II: First Half vs. Second Half	1008: − 46	1008: −38	.10
Series I vs. Series II	2016: + 38	2016: −84	1.05

Series II did not show any significant change in scoring between its first and second halves.

When Series I and Series II are compared to see whether there was any chronological trend in scoring over the entire period covered by the research, we find that the scoring declined from +38 in the first half of the research (Series I) to −84 in the second half (Series II). The difference here is 122 hits. The standard deviation of the difference is 115.85 and the CR, 1.05, which is statistically insignificant.

DISCUSSION

From the point of view of the primary objective of this research, that of locating individuals with outstanding PK ability, the investigation was not highly successful. Only one subject, whose results have already been described in an earlier report (2), showed outstanding individual performance in the tests. It has seemed worth while, however, to report all the results here for two reasons: first, the experiment was done under sufficiently safeguarded conditions to make the data of scientific interest in relation to the PK hypothesis, regardless of the total results. Second, the analyses made for position effects revealed scoring trends which were to be expected on the basis of numerous earlier investigations.

There is, of course, some question regarding the proper statistical significance attaching to these findings in view of the fact that no one anticipated that all of these analyses would be made at the time the research was planned. At the same time, it will be evident to any student of the research that the analyses for position effects are those which have become customary for all PK investigations,

and it will also be evident that, to the extent that the analyses yielded suggestive results, the differences observed are in line with the findings of previous studies.

The most outstanding feature of the results was the U-curve shown in the vertical distribution of Series I. This finding parallels closely the vertical distribution found in the work of Reeves (5) and in an earlier series by the present investigator, in which both the beginning and the ending trials of the column gave higher scores than the middle trials. Other reports have shown the highest scores in the first trials of the first column only, particularly in the first column of the set (**1, 3, 4, 5, 6, 7**).

Many investigators have observed the outstanding rate of scoring in the very first trial of the set. It was decided, therefore, to analyze the present data for this same effect by isolating the first trial of each column, which in this work was the equivalent of a set. The data of this trial position comprised 168 runs in each series. In Series I, the deviation obtained was +65, and in Series II, +38. Pooling these results for the entire investigation gives 336 runs with a deviation of +103. The SD is 33.46 and the CR, 3.08 ($P = .001$). This finding has suggestive value only, in this study, but the frequency with which this strong beginning salience in PK has been observed indicates that it is a fairly general effect.

The other outstanding trend in these results is the decline between the first and second halves of Series I ($CR_d = 2.76$). Assuming for the moment that this decline is not attributable to chance variation, it is not possible to assign it to a single factor in the test conditions. There is a possibility, of course, that it is a genuine chronological decline which set in as the novelty of the tests wore off. But another factor which may have contributed to it is the shift from one to another order of throwing for the target faces at the midpoint of the series. The subjects had got accustomed to the *ascending* order of targets, and the change to the descending order might have upset their scoring. Perhaps the latter interpretation is favored somewhat by the fact that earlier analyses of PK data for the chronological distribution of successes in cases where the experimental conditions have remained constant have not as a rule revealed any significant change in scoring.

The data of the present research offer general confirmation of the fact that the PK process is influenced in subtle psychological ways by the structure of the task presented to the subject. At the same time, these results suggest that the occurrence of PK in a test may be overlooked if the total score is the only criterion of its effect which is used.

References

1. Dale, L. A. The psychokinetic effect: the first A.S.P.R. experiment. *J. Amer. Soc. psych. Res.*, 1946, **40**, 123-51.
2. Gibson, E. P. Note on an impromptu experiment in psychokinesis. *J. Amer. Soc. psych. Res.*, 1947, **41**, 22-28.
3. Herter, C. J., and Rhine, J. B. An exploratory investigation of the PK effect. *J. Parapsychol.*, 1945, **9**, 17-25.
4. Pratt, J. G. Lawfulness of position effects in the Gibson Cup PK Series. *J. Parapsychol.*, 1946, **10**, 243-68.
5. ———. Restricted areas of success in PK tests. *J. Parapsychol.*, 1947, **11**, 191-207.
6. ———. Rhythms of success in PK test data. *J. Parapsychol.*, 1947, **11**, 90-110.
7. Pratt, J. G., and Woodruff, J. L. An exploratory investigation of PK position effects. *J. Parapsychol.*, 1946, **10**, 197-207.
8. Rhine, J. B., and Humphrey, B. M. Special evidence from hit patterns: I. Quarter distributions of the page. *J. Parapsychol.*, 1944, **8**, 18-60.

1221 Philadelphia Avenue, S.E.
Grand Rapids, Michigan

BOOK REVIEWS

EMANUEL SWEDENBORG, by Signe Toksvig. New Haven: Yale University Press, 1948. Pp. 389.

This book is a mine of information about the life, personality, scientific and practical achievements, and theological doctrines of a very remarkable man who is too little known to those of us who are not members of the church which is founded upon his alleged revelations. My impression is that Swedenborg has been the object of greater attention and admiration in England and the United States than in his Swedish fatherland; but that is not saying much.

His family name was originally Swedberg, and he was born in Stockholm in 1688 to Jesper Swedberg and his first wife, Sara Behm. The father was an almost pathologically self-satisfied cleric who was highly successful both in his profession and in his three marriages. After filling important offices in the University of Upsala, he became Bishop of Skara in 1702 and died in the enjoyment of that preferment in 1735 at the age of eighty. In 1719 his children were ennobled and took the more distinguished name of Swedenborg.

Emanuel became a student at Upsala University in 1699 at the age of eleven and took his degree there in 1709. At that time and for many years afterwards he was passionately interested in mathematics and natural science, both theoretical and practical. In the following year he went to England and stayed there pursuing his studies for himself until 1715. Soon after his return he acted as assistant to the distinguished Swedish engineer, Polhem. On Polhem's advice Charles XII appointed him to an assistantship in the Board of Mines in 1717. This Board was an important department of state which exercised complete supervision over Sweden's ancient and already very extensive iron and copper mining industry. Swedenborg remained an active officer of the Board, with intervals of leave of absence abroad granted to him for purposes of study and publication, until 1747. In the spring of that year the important post of Counsellor of Mines fell vacant; the Board unanimously recom-

mended Swedenborg for the job; and the king offered it to him. He declined, and instead asked and obtained leave to retire on half-pay. He was then sixty years old; he was a comparatively wealthy man of simple tastes; and he believed that it had been revealed to him that his life work was to provide a symbolical interpretation of the Bible under divine instruction.

The author makes it abundantly clear that Swedenborg was a first-rate scientist, engineer, administrator, and member of the House of Nobles. There was nothing "cranky" about his behavior, either in social life or in the highly technical and responsible official business which he performed to the general satisfaction of his colleagues and his employers. Apart from this, it is obvious that a man who could, up to the age of eighty, travel and sojourn abroad to the extent to which Swedenborg did, without any personal attendant, at a time when there were no travel agencies or advance bookings by wire, must have had his head screwed on the right way in practical life. His main contributions to pure science were in theoretical astronomy and in the anatomy of the brain and nervous system. The author produces testimony from experts that they were of the highest originality and importance.

We can now pass to the peculiar experiences which played so great a part in the inner life of this outwardly normal and sober-minded individual. There is a tradition that Swedenborg, like many others, claimed as a young child to have invisible playmates. But there is no trace of anything psychic or mystical or even particularly religious in anything that is known of him from that time until about 1736 when he described experiences of light and brightness and flame which he had repeatedly had and which he took to be a confirmatory sign of the correctness of his intellectual processes.

In 1743 he had a series of mystical experiences of a specifically Christian kind which culminated at Delft on April 6, 1744, with an experience of direct personal intercourse with Christ. By 1745 he had become convinced that his life work was to interpret the scriptures symbolically under the direct instruction of their author. Soon after April, 1745, and thereafter for the rest of his life, he repeatedly had experiences in which he seemed to himself to see and hear and touch and converse with the souls of the dead in their own

environment while he himself was awake and in an otherwise normal state of consciousness.

Part of Swedenborg's revelations came to him in the form of automatic script. The author has inspected much of this which is preserved in Stockholm and has noted very different handwriting in different scripts produced at much the same time. Samples are reproduced in the book. The author relates that when photographs were submitted without comment to a handwriting expert, the latter assigned them to different writers. There is some correlation between the style of the script and the character of the contents. There is, for example, a violently contorted kind of script which is concerned mainly with four topics, of which the most notable is violent polemics against certain characters in the Old Testament.

Swedenborg gives valuable accounts of his experiences as they presented themselves to his introspection. He states also that he had discovered in childhood while at prayer, and had developed and practiced ever since, certain methods of breathing which induced mental concentration and favored intercourse with the spiritual world. The author points out the analogy to the traditional methods of Indian yogins.

On the basis of what he seemed to himself to perceive and to be told in the spirit world, Swedenborg describes in great detail the nature and laws of life in it. Much that he learned came as a great surprise to him. He had believed, on general philosophical grounds, that the next life must be purely immaterial and intellectual. He found to his surprise that the souls of the departed continue to have something analogous to bodies and something analogous to sensation. In consequence of this, newcomers are liable to believe for some time that they still have ordinary material bodies. They sometimes cannot be persuaded that they have died. Again, mutual communication between departed spirits and communication between a departed spirit and an embodied one, such as Swedenborg, in fact takes place by a telepathic conveyance of symbolical visual images. But these are automatically translated into verbal images in the language which the recipient habitually speaks. As a result, spirits are liable to believe that they literally talk and communicate by speech. Swedenborg had many acute arguments with spirits to prove to them that they were mistaken in such beliefs.

From the account which the author gives of Swedenborg's reports on the world of spirits, I should be inclined to summarize the doctrine very roughly as follows: Each spirit lives in a kind of dream world created by autosuggestion and telepathy and expressive of his beliefs, emotions, desires, and character. Since there is constant telepathic interaction, the dream worlds of associated spirits cohere with each other; and so, for practical purposes, such spirits live in a common environment, though it is considerably more plastic than that of men in their present life. In this way spirits create their own hells and heavens; the scenery and the contents which they ostensibly perceive are mind-created, but the pains and the pleasures which they feel are real, as they are in our dreams. The hells are inhabited by self-centered egoists, stripped of the masks which social constraint and self-deception imposed on their real desires and emotions during their earthly life. The heavens are inhabited by spirits who take a direct loving interest in other persons and things for their own selves and without regard to the reactions of the latter upon themselves. In the spiritual world, like gravitates toward like; the cruel and selfish torment each other, and the gentle and unselfish minister to each other's well-being. In the present life mental and even material conditions are greatly influenced by the spirits whom we attract into our sphere. From time to time conditions in the spiritual world degenerate greatly with consequent bad effects on the earthly plane. When things reach a certain degree of badness God holds a kind of spiritual spring cleaning. Swedenborg believed that one of these "judgments" had taken place in 1757 and that he had witnessed it. It was the *last* judgment only in the sense of the *latest*.

Obviously most of us have no means of applying any objective tests to these statements. At the best we can regard them as "travellers' tales" and can note, for what they are worth, agreements and disagreements between them and statements made on the same topics by other mediums or mystics. We must remember, too, that, while there is nothing specially absurd or implausible in these statements, they are mixed up with fantastic allegorical interpretations of the scriptures and with the doctrine that the material universe has the form of a vast human body.

On the other hand, there are certain stories of paranormal cognition on Swedenborg's part, to which we can apply the ordinary tests used by critical psychical researchers. The author quotes those main cases and those less well attested ones. The main cases are (1) Swedenborg's alleged clairvoyant awareness on a certain day in July, 1759, of a contemporary serious fire in Stockholm when he was himself at Güteburg about 460 kilometers away. (2) His alleged revelation to Queen Louisa Ulrica of Sweden of the gist of her last (and secret) letter to her brother, the Prince of Prussia, who had died before he could answer it. (3) His alleged revelation to the widow of the Dutch ambassador in Stockholm of the precise place, in an unsuspected secret compartment of a certain bureau belonging to her late husband, in which she would find the lost receipt for a service of plate bought and paid for by him before his death. All these stories are mentioned by the philosopher, Kant, who was interested in Swedenborg's *alleged* feats. The main authority for the first and the third is Kant's friend, the English merchant, Green, whom he commissioned to investigate the evidence when in Sweden. For the second we have now better evidence than any available to Kant; namely, that of a contemporary diary of Count Tessin and another narrative written after the Queen's death by Count von Höpken which throws light on the very peculiar circumstances under which the correspondence between the Queen and her brother took place.

I should say that the evidence in none of these cases reaches the very high level attained by the best-attested stories in the archives of the Society for Psychical Research. But those which I have quoted are quite good second-class cases; and, if it be granted (as it must be) that there is adequate evidence for the occurrence of paranormal cognition, they suffice to make it quite likely that Swedenborg possessed that gift to a high degree.

Various psychoanalytic writers on the Continent have pronounced on Swedenborg's case from the standpoint of psychopathology with that wealth of detailed assertion and that paucity of empirical evidence which is the "badge of all their tribe." The author quotes some entertaining examples. Swedenborg has been found by one of these experts to be the victim of an "unsolved inverted Oedipus

complex due to a repressed homosexual attachment to his father," the bishop. I think we can afford to dismiss this stuff with a smile or a sigh.

Swedenborg enjoyed extremely good bodily health until his eighty-third year when he had a stroke of paralysis, from which he never fully recovered, while on a visit to London. He was devotedly nursed by Mr. and Mrs. Shearsmith, at whose home he lodged, and by their maid, Elizabeth Reynolds. He died in the presence of Mrs. Shearsmith and Elizabeth very peacefully on Sunday, March 29, 1772. He was buried in the Swedish church in London, and his body rested there until 1908, when it was taken back to his fatherland by the Royal Swedish battleship, *Fylgja*.

I am in no position to check the accuracy of the author's statements about Swedenborg's life and experiences and doctrines, but I have no reason to doubt that they are substantially correct. Certainly this is a very interesting book about a very remarkable man, and it will repay the attention of psychologists, psychical researchers, philosophers, and theologians. I hope I shall not be accused of what Swedenborg regarded as the root of all evil, namely, egoism, if I point out that it contains two statements about myself which are severally false and collectively inconsistent. On page 159 I am described as "Professor C. D. Broad of *Oxford*" and on page 168 as "Professor of *Logic* at Cambridge, C. D. Broad." I have italicized in each case the part of the statement which is incorrect.

Finally I would add that the author has the advantage of being familiar with and sympathetic to the recent developments of psychical research. In particular, the help of Professor Gardner Murphy and the late Mr. Whately Carington is gratefully acknowledged. I should consider that such knowledge is an absolutely essential prerequisite to any satisfactory study of Swedenborg's life and work, and I should very much doubt whether it has been possessed by any other writer who has treated the subject.

C. D. Broad

Professor Broad is Fellow of Trinity College and Knightsbridge Professor of Moral Philosophy at the University of Cambridge.

LA RICERCA PSICHICA [Psychical Research], by Emilio Servadio. Preface by Charles Richet. Rome, 1946. Pp. 142.

This is the second edition of a book first published in 1930 by a well-known Italian psychoanalyst. Servadio is one of the few medical men who combine a sound psychiatric background with a comprehensive knowledge of the literature of parapsychology. This puts him into a position to survey the field from an excellent vantage-point, although the reader may wish he had devoted more space to psychoanalytic aspects. Even so, his slim volume covers a remarkably wide range. A brief historic account is followed by a well-balanced discussion of mediumship and its various manifestations. The experimental approach developed by the "Durham School" is given due consideration, including more recent results of PK experiments. The Carington and Soal-Goldney tests are also referred to in passing. No attempt at a critical appreciation of the material or of existing theories is made, although the author seems to lean towards Freud's views about the archaic, regressional nature of telepathy and related phenomena, and to Geley's and Driesch's neo-vitalistic interpretations.

Servadio concludes in quoting Richet's remark that psychical research, though it cannot as yet be considered a science in its own right, is the great hope for the future: the science of tomorrow. Richet's foreword to the first edition, written eighteen years ago, contains an eloquent appeal to the skeptic to read the book and to acquaint himself with the facts before rejecting them without a hearing. His appeal is still up to date.

JAN EHRENWALD, M.D.

Dr. Ehrenwald is a practicing psychiatrist in New York City. He is a Fellow of the Royal Society of Medicine and an Associate in Psychiatry at the Long Island College of Medicine.

LETTERS AND COMMENTS

To the Editor:

I am greatly interested in your experiments and am certain that collaboration on a research project would be mutually productive. Spontaneity tests and simultaneous group tests of all ESP subjects should be made in order to determine whether any correlation exists between ESP productivity on one hand, and the subjects' spontaneity quotients and sociometric position in the group on the other hand. It appears to me that the focal interest should be given to the increase or decrease of ESP productivity, as compared with the increase or decrease of the spontaneity quotient and change in sociometric status.

In order to obtain such information the group of ESP subjects has to be studied over a considerable period of time, at least six months or one year. The spontaneity tests and group tests should be repeated at weekly or bi-weekly intervals. A control group might be set up in conjunction.

A second project of interest to you might be the following: to get two subjects of about equal ESP productivity in the past, train the spontaneity of one in situations which do not necessarily relate to ESP but which involve sensitivity of perception, and to leave the spontaneity of the other subject untrained. After this training period their ESP productivity should be tested again. Then we would have an opportunity to test the hypothesis that both the "tele" factor and the "spontaneity" factor operate in the sphere of ESP, as they have been found to operate within the sphere of social relations and within the sphere of action respectively.

J. L. Moreno, M.D.

Sociometric Institute, Inc.
101 Park Avenue
New York, N. Y.

TO THE EDITOR:

It was with profound interest and appreciation that I studied the article prepared by Dr. J. B. Rhine, appearing in the March issue of the JOURNAL, concerning the conditions which he had found best favored success in psi tests.

It caused me to review the great advance made in methods of ESP research and experimentation during the past ten years and to reflect upon my personal experiences with this field of phenomena in the "long distance telepathic tests" conducted with Sir Hubert Wilkins in the Arctic as sender, and myself in New York as receiver.

It had been Wilkins' plan to get off by himself at the appointed times and review in his mind the important happenings of the day, *willing* them consciously to me. But, because of the pressure of his many rescue duties, he often was unable to keep such appointments. He kept, however, a diary and log of events and formed the practice of thinking strongly of me at different odd moments during the day. I was amazed to discover that the accuracy of my impressions did not seem to have been impaired on the occasions when Wilkins was not able to "send" during my "sittings."

These faculties of inner consciousness apparently operated for me by reaching out in time and space and bringing back an awareness of certain conditions or circumstances or thoughts in the minds of others, utilizing whatever phase of ESP was required at the moment to ascertain this knowledge.

But now we have learned, as Dr. Rhine points out, that the personality and approach of the investigator is as great a contributing factor to the success of an experiment as the developed powers of the subject himself.

I am glad to see a growing concurrence among workers in the ESP field that the generation of an attitude of interest, enthusiasm, and confidence in the subject toward his tests is productive of more effective results.

Because of my temperament, I am much more intrigued with tests having to do with real life experience. If the sender is reacting emotionally to something he is doing which is not routine, his own interest in the experiments is kept at a high pitch and he is auto-

matically generating real feeling behind his thoughts, affording the subject a much greater opportunity to receive impressions of a telepathic, clairvoyant, or precognitive nature. I should like to see a series of controlled "real life tests" worked out by investigators.

Dr. Rhine has emphasized the value of letting the subject know, at the end of the tests, how well he has been doing, following the "party or game" plan of procedure and atmosphere. I wish to confirm the wisdom of such a practice on the part of the investigator as it influences the effectiveness of the subject. In my own experience, there was a period of almost two and a half months when I recorded impressions, three nights a week, without receiving any confirming word from Wilkins, who was beyond the reach of airmail most of this time. Under these circumstances, it was increasingly difficult for me to keep up my enthusiasm and maintain confidence in the impressions I was recording.

We lack words, as yet, to describe much of what happens in the mind of the subject demonstrating psi, but the day of real achievement is coming—even though our steps may be faulty, our methods yet crude, and our progress slow.

Sincerely yours,

HAROLD M. SHERMAN

Mountain View
Arkansas

DEAR DR. RHINE:

I should like to make some comments on the review of my book, *Spiritism or Parapsychology?*, reviewed in the JOURNAL OF PARAPSYCHOLOGY for June, 1948:

1. Dr. Ehrenwald [the reviewer] probably did not understand my book entirely since it was written in the Slovak language, which is similar to but not identical with the Czech.

2. I am Professor of Theology in the University of Bratislava, not of Trnava, the second largest town of Slovakia.

3. The critical evaluation of parapsychological facts is given on page 151-200 where I discuss the different theories. The whole book is a critical evaluation of the problem: spiritism or parapsychology? as is indicated in the title.

4. The following viewpoint is presented in the interpretation of the facts: the mind communicates directly with mind (p. 161) and acts directly on things (p. 164). These conclusions are in agreement with those reached in experiments at Duke University and elsewhere.

5. The author does take "a definite stand in discriminating between the scientific approach of parapsychology and the emotionally colored beliefs of the spiritualists." Spiritism in general is refuted on pages 126-151. The existence of spirits is admitted only in exceptional cases (for example, Mirabelli). The author, being a Catholic priest, cannot entirely deny the existence of spirits, and there are certain facts which testify to their existence.

6. "The anecdotal accounts given to the author by the country folk" are not anecdotes but facts and are well distinguished from parapsychological facts (pp. 46-51). The author also evaluates them from the point of view of hysteria. Therefore the student can very well distinguish in the book between science and purely naïve folklore.

Yours faithfully,

ALEXANDER SPESZ, M.D.

Kapitulská
Bratislava, Czechoslovakia

TO THE EDITOR:

Responsive to the invitation that interested persons contribute their views as to what lines of research should be followed in parapsychology during the next decade, I submit herewith several random suggestions.

First of all, I would emphasize Dr. Murphy's viewpoint on the treatment of spontaneous cases. May we not be allowing significant clues to slip through our fingers simply through lack of anticipatory planning? The striking instances that come to mind, cases often seized upon by sensational journalism but not immediately, if ever, subjected to investigation by competent researchers, must be matters of keen regret to everyone who appreciates the enormous difficulties under which parapsychology, from its sheer dearth of knowledge of basic principles, necessarily labors.

I would suggest also that experiments already initiated be extended to their utmost ramifications, not merely by a few researchers, but as widely as possible. Here, it seems to me, the impressive results obtained by Dr. Rhine, as compared with the poor or negative results of British researchers, provide an object lesson. Positive results by one experimenter in parapsychology do not invariably presuppose comparable results by another; but the reverse should also be true. If so much as a single experiment eventuates suggestively, subsequent negative results should not convince us too readily that the original experiment was invalid: the particular phase of research should be pursued exhaustively by many researchers at different times and in various places.

It seems scarcely necessary to plead for open-mindedness on the part of researchers in parapsychology. It may be relevant, however, to raise the question as to just how far, in actual practice, they are prepared to go against accepted theories of firmly established branches of science. One occasionally runs across reports of occurrences which by prevailing tenets are fantastic, but which, on the other hand, are challenging to what might be termed "the parapsychological attitude." Investigation would seem to be in order, regardless of the matter of running counter to authority and even to what we are accustomed to think of as common sense.

Very truly yours,

F. Frederick Clouser

New Bloomfield, Pennsylvania

To the Editor:

After digesting the tenor of the March issue of your journal, I thought I would comment on a few points. A large part of the issue represented a plea (in a sense a prayer) for the repeatable experiment.

In this respect the writers of the articles appear to run counter to statements made in earlier issues. It is clear that extrachance results in one experiment are not nullified by the failure to obtain similar results on a repeat.

Now, of course some British experimenters (as one British

author states) are puzzled at not being able to confirm all the results produced by the American researchers. This situation naturally has engendered difficulty for the British parapsychologists.

Is this mathematically sound? Certainly it is psychologically human, but is it not inconsistent with scientific objectivity?

Now, if an American experimenter obtains a P = .0001 for the PK effect, and if his work is repeated by five separate research groups in England with no extrachance results, we must conclude (arbitrarily weighing these experiments equally) that for the PK effect the scientifically established probability is P = .0005. I recall that in one issue of the JOURNAL there was a master summary of all the experiments bearing upon a certain ESP effect, with P equal approximately to 10^{-6}. On the basis of this type of reasoning, it has been concluded that further experiments to verify clairvoyance, precognition, and PK are unnecessary. This argument and position are unassailable. Why does acceptance fail to follow even among sympathetic academicians?

Naturally it is true that repeatable experiments are desirable, but there is no reason why they should be obtainable (except with a few sensitives). Laboring to set up conditions tending to make experiments repeatable will increase parapsychology's prestige, but the real situation appeared distorted in the symposium.

In this respect, I would like to mention an experiment I conducted with my mother (a sensitive). A randomized deck of regular playing cards was presented (at a short distance), and the subject guessed the suit of the first 28. (The deck was untouched.) The results were then tabulated. There were 17 correct calls as against 7 by chance, and the probability obtained was P = .000049. No further experiments were run with this subject. Reasons: the result was markedly significant, and nothing new was to be gained since the conditions are not adequate for throwing further light on the psi process at work. For example, they do not distinguish between clairvoyance and precognitive telepathy. A new procedure would have to be developed before a further series of experiments would be desirable.

As a physicist, I read with interest the symposium speculations on the PK force. One author mentions psychopods and similar

effects produced in spiritual séances, and implies their connection with the PK force. Rank speculation! In closer agreement with known results is the concept that, in keeping with its mental origin, the PK force is mental in nature. It has so many nonphysical attributes as to demand the conclusion that no materialized instrumentality will fulfill the requisites.

As for the physicist's approach to, and mathematical concept of, this force (the Ψ_κ of the British), I have yet to analyze the possibilities. Concerning the experiments suggested, the theoretical possibility of obtaining positive results is existent, but the variability of results and the low probability of success argue for greater care in the construction of the experimental procedure.

A final point of interest is the reference to astral projection, etc. I've studied these occult phenomena very closely. Not statistical studies, but case histories suggest their existence, and these matters certainly are amenable to scientific investigation. Personally I have experimented in occultism: (1) I have induced astral projection, but not to the extent that I can subject it to statistical and quantitative analysis by others; (2) I have statistically verified Dunne's "experiment with time."

Sincerely yours,

DOUGLAS STEEN

807 West Forty-first Street
Los Angeles 37, California

GLOSSARY

In order to avoid constant redefining of commonly recurring terms in papers appearing in this JOURNAL, the following definitions are submitted for convenient reference.

A simple description of the basic experimental methods and of the statistical procedures to be followed in evaluating ESP (both card tests and drawings tests) and PK may be found in the *Handbook of Tests in Parapsychology* by Betty M. Humphrey (published by the Parapsychology Laboratory, Duke University, Durham, N. C.).

AGENT: In tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

AVERAGE SCORE: Average number of hits per run.

CALL v.: To attempt to identify by ESP a stimulus object or a mental state of an agent.

CALL n.: The response described above; also the resulting selection.

CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

CHANCE EXPECTATION = MEAN CHANCE EXPECTATION: The most likely score if only chance obtains.

CHANCE AVERAGE: Mean chance expectation expressed as an expected score generally in terms of average per run.

CHI-SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic perception of the mental state of another person.

CR (CRITICAL RATIO): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearson's.)
CR OF THE DIFFERENCE: The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

DECK: Twenty-five ESP cards, five of each suit.

DEVIATION: The amount an observed number of hits or an average score varies from the mean chance expectation or chance average. A deviation may be total (for a series of runs) or average (per run).

DIE THROW: The elementary unit of data in a dice-throwing test represented by the throwing and reading of a single die, no matter how many dice are thrown at a time.

EMPIRICAL CONTROL: An experiment which wholly or partially follows the main experiment with the exception that the conditions are designed to exclude the possibility of the hypothesis being tested.

ESP (EXTRASENSORY PERCEPTION): Response to an external event not presented to any known sense.

ESP CARDS: Cards, each bearing one of the following five symbols: star, circle, three parallel wavy lines (called "waves"), square, plus.

ESP SYMBOLS: See plate opposite page 1, this JOURNAL, Vol. 1, March, 1937.

EXPECTATION; see CHANCE.

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

FREE RESPONSE METHOD: An ESP test method in which a relatively unlimited range of stimulus objects is used and in which the subject is instructed to express freely as his response any range of thoughts and ideas.

GESP (GENERAL EXTRASENSORY PERCEPTION): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

HIGH-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling eight or more.

LOW-DICE TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling six or less.

MEAN CHANCE EXPECTATION; see CHANCE.

P (PROBABILITY): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with those psychical effects which appear not to fall within the scope of what is at present recognized law.

PERCIPIENT: The person who makes the calls in an ESP test.

PK (PSYCHOKINESIS): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring free responses. A judge ranks the stimulus objects (usually in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

PSI: A general term to identify personal factors or processes in nature which transcend accepted laws. It approximates the popular use of the word "psychic" and the technical one, "parapsychical."

PSI PHENOMENA: Occurrences which result from the operation of psi. They include the phenomena of both ESP (including precognition) and PK.

QD (QUARTER DISTRIBUTION): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in one run.

TOTAL SCORE: Total of scores made in a given number of runs.

AVERAGE SCORE: Total score divided by number of runs.

SD (STANDARD DEVIATION): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$, in which n is the number of single trials, p the probability of success per trial, and q the probability of failure. (For ESP cards, SD $= 2 \sqrt{\text{no. of runs.}}$) Sometimes the SD used is that estimated from the observed variability in the scores.

SD OF THE DIFFERENCE: For both ESP cards and PK tests using dice, the SD of the difference is equal to $\sigma_s \sqrt{1/R_1 + 1/R_2}$ where σ_s is the SD of a single run and R_1 and R_2 are the number of the runs in the respective samples compared. This gives the SD of the difference for run score averages.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SEVENS TEST: A PK technique in which the aim of the subject is to try to influence a pair of dice to fall with the two upper faces totaling seven.

SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. Common criteria are: a probability value of .01 or less, or a deviation in the expected direction such that the critical ratio is 2.33 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

TARGET: In ESP tests, the stimulus object; or in telepathy, the mental state of the agent. In PK tests, the faces of the die (or combination of faces) which the subject attempts to bring up in the act of throwing.

TARGET CARD: The card which the percipient is attempting to perceive (i.e., to identify or otherwise indicate a knowledge of).

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET DECK: The deck of cards the order of which the subject is attempting to identify.

TARGET FACE: The face on the die which the subject tries to turn up as a consequence of direct mental action.

TELEPATHY: Extrasensory perception of the mental activities of another person. It does not include the clairvoyant perception of objective events.

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.

JOURNAL OF PARAPSYCHOLOGY

VOL. 13 , 1949
NOs 1-4

PUB. 33

UNIVERSITY MICROFILMS
ANN ARBOR, MICHIGAN. 1950

THIS PUBLICATION IS REPRO-
DUCED BY AGREEMENT WITH THE
COPYRIGHT OWNER. EXTENSIVE
DUPLICATION OR RESALE WITH-
OUT PERMISSION IS PROHIBITED.